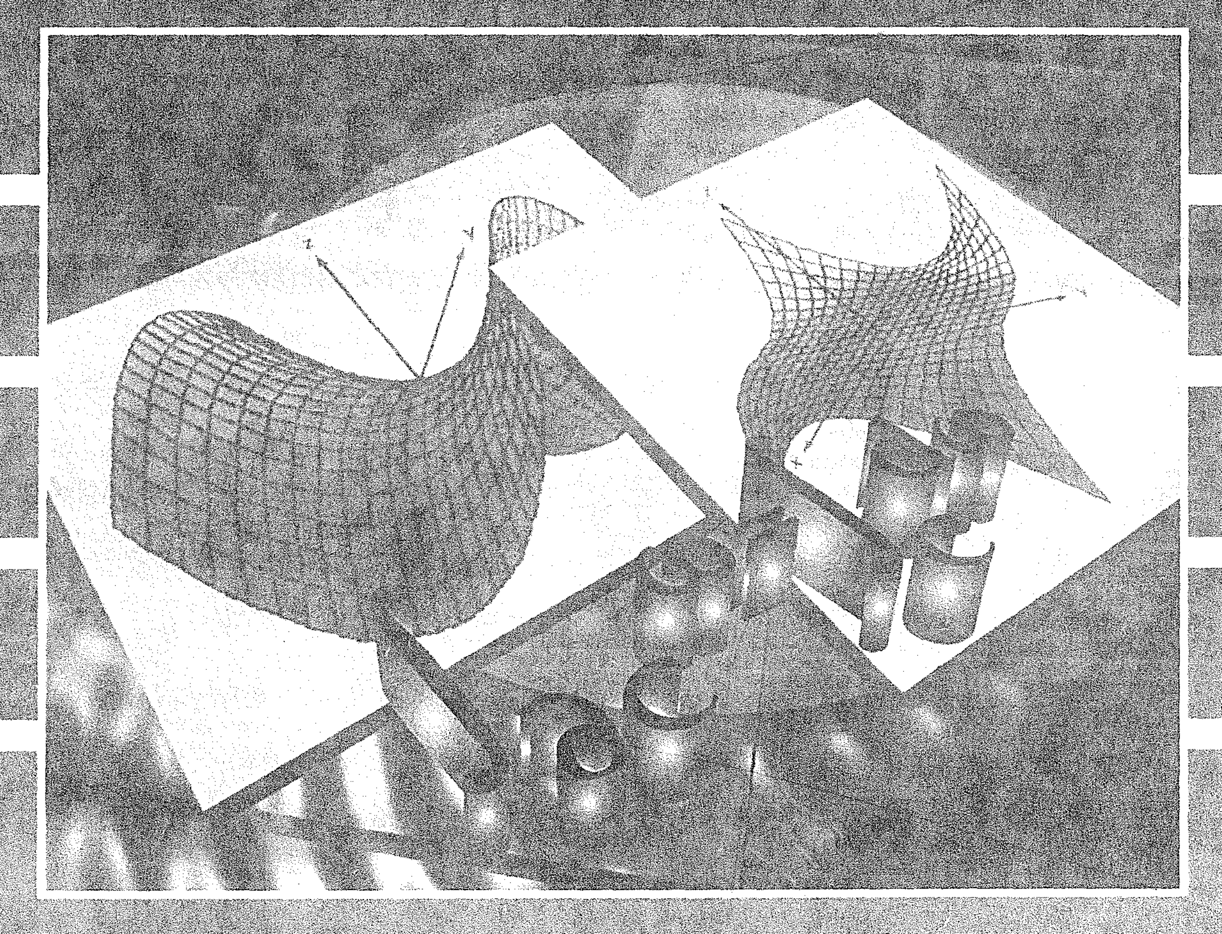
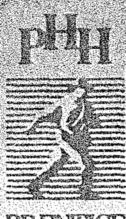
CALCULO VECTORIAL

Claudio Pita Ruiz





PRENTICE-HALL HISPANOAMERICANA, S. A.

Cálculo Vectorial

PRIMERA EDICIÓN

Claudio Pita Ruiz

Universidad Panamericana
Escuela de Ingeniería



PRENTICE
HALL

MÉXICO • NUEVA YORK • BOGOTÁ • LONDRES • MADRID
MUNICH • NUEVA DELHI • PARÍS • RÍO DE JANEIRO
SINGAPUR • SYDNEY • TOKIO • TORONTO • ZURICH

EDITOR: Luis Gerardo Cedeño Plascencia
SUPERVISOR DE TRADUCCIÓN: Jorge Bonilla Talavera
SUPERVISIÓN PRODUCCIÓN: Julián Escamilla Liquidano

Pita: Cálculo Vectorial 1/Ed.

Calle 4 Nº 25-2º piso Fracc. Ind. Alce Blanco,
Naucalpan de Juárez, Edo. de México,
C.P. 53370

ISBN 968-880-529-7

Miembro de la Cámara Nacional de la Industria Editorial, Reg. Num. 1524

SEP

PROGRAMAS EDUCATIVOS, S. A. DE C.V.
CALZ. CHABACANO No. 65, LOCAL A
COL. ASTURIAS,DELEG. CUAUHTEMOC,
C.P. 06850, MÉXICO, D.F.

EMPRESA CERTIFICADA POR EL
INSTITUTO MEXICANO DE NORMALIZACIÓN
Y CERTIFICACIÓN A.C., BAJO LA NORMA
ISO-9002: 1994/NMX-CC-004: 1995
CON EL No. DE REGISTRO RSC-048

2000 1998

It seems to be one of the fundamental features of nature that fundamental physics laws are described in terms of a mathematical theory of great beauty and power, needing quite a high standard of mathematics for one understand it. You may wonder: why is nature constructed along these lines? One can only answer that our present knowledge seems to show that nature is so constructed. We simply have to accept it. One could perhaps describe the situation by saying that God is a mathematician of a very high order, and He used very advanced mathematics in constructing the Universe.

Paul Dirac

Let us grant that the pursuit of mathematics is a divine madness of the human spirit.

Alfred North Whitehead

Prólogo

The values [of mathematics] are there, values at least as great as any human creation can offer. If all are not readily or widely perceptible or appreciated, fortunately they are utilized. If the climb to reach them is more ardous than in music, say, the rewards are richer, for they include almost all the intellectual, aesthetic, and emotional values that any human creation can offer.

Morris Kline

Este es un libro de cálculo diferencial e integral de funciones cuyo dominio y/o codominio son subconjuntos del espacio $\mathbb{R}^n$. Como a los elementos de este espacio se les llama "vectores", un nombre popular para este tipo de temas dentro del cálculo es el de "cálculo vectorial". De otro modo aún, este libro trata sobre el cálculo en (espacios de) dimensiones superiores. El único prerrequisito formal para estudiar el material que aquí se presenta, es haber tomado un curso de cálculo diferencial e integral de funciones reales de una variable real (como el que se estudia en un primer semestre de cálculo), junto con algunos resultados elementales sobre sistemas de ecuaciones lineales y matrices (que se estudian generalmente en un curso de álgebra superior o en los primeros capítulos de un curso de álgebra lineal).

El cálculo es el primer contacto de un estudiante con la llamada "matemática superior"; desde el concepto de límite para funciones de una variable se puede advertir que las ideas que se manejan en esta parte de la matemática tienen un sabor diferente de las que se habían estudiado previamente (álgebra, trigonometría, geometría analítica). Actualmente ya no es necesario insistir en la importancia del estudio del cálculo, como primera etapa para adentrarse en problemas matemáticos más elaborados, o bien para abordar problemas en otras ramas del conocimiento que utilizan de manera importante las herramientas que ofrece el cálculo. Esta parte de la matemática fue, desde su nacimiento en el siglo XVII, es ahora, y seguirá siendo, la antesala de los problemas propios del estudio de la mayor parte del conocimiento científico actual, como el que aparece en los planes de estudio de las carreras de ingeniería o ciencias. Esto es especialmente cierto con los temas del cálculo en dimensiones superiores, como los que contempla este libro. Lo es, por ejemplo, por las importantes aplicaciones que de estos temas se derivan, sobre las cuales puse una especial atención para que aparecieran en los momentos importantes del desarrollo de la teoría. Por otra parte, el cálculo en dimensiones superiores nos brinda la primera oportunidad de disfrutar las satisfacciones intelectuales que proporcionan los procesos de generalización en matemáticas. Una vez entendidos los conceptos del cálculo para funciones reales de una variable, y que se admira la fuerza de estas ideas para resolver problemas en otras partes del conocimiento científico, aún más, cuando llegamos a pensar que estamos pisando terrenos "muy elevados" de la matemática, el cálculo en dimensiones superiores nos muestra que estábamos apenas a la mitad de la montaña, y que las emociones fuertes apenas comienzan a aparecer al ver que los resultados del primer curso de cálculo son casos particulares de situaciones que contemplan los mismos problemas, pero de una manera más general.

Esta obra contiene más material del que se puede cubrir normalmente en un segundo curso de cálculo con estos temas. No es, sin embargo, un tratamiento exhaustivo del cálculo en $\mathbb{R}^n$. Como en cualquier libro de matemáticas, hay varias ausencias (por ejemplo, las demostraciones de los teoremas

de la función implícita y de la función inversa que se estudian en el capítulo 3), y la justificación de estas ausencias es también, como en cualquier libro de matemáticas, la misma: no es posible tener en unas cuantas páginas todos los temas que contempla y que se derivan de una (cualquiera) parte de la matemática. Los temas tratados en los libros de matemáticas son fruto principalmente de dos motivaciones del autor. La primera de ellas es que el libro debe contener como mínimo el material que se debe cubrir en un curso normal. La segunda es que el libro debe ofrecer más que este material mínimo (de otra forma se podría convertir en una recopilación de apuntes del curso), ya sea profundizando en los temas tratados, o bien, presentando algunas de sus derivaciones. Y son los gustos y las debilidades matemáticas del autor los que deciden el producto de esta segunda motivación, lo cual provoca entonces la ausencia de algunos temas, así como el estudio de algunos temas no usuales en un curso sobre la materia. Lo que presentamos en este libro se no es ajeno a estos hechos, pues éste contiene como subconjunto propio el material "normal" de un segundo curso de cálculo . . . y algunas cosas más. Las partes correspondientes al complemento de los temas obligados en un curso de esta materia, que considero son las "más prescindibles" en un primer acercamiento al cálculo en $\mathbb{R}^n$, aparecen como apéndices de secciones de capítulos, o bien como secciones que están marcadas con un asterisco. Con estas indicaciones explícitas, y el criterio (y gusto) del profesor, se pueden planear varios programas de cursos en los que se puede usar el presente libro como texto.

El inicio de esta obra "considera" el conjunto $\mathbb{R}^n$, formado por n-adas ordenadas de números reales, y termina con la demostración del teorema (general) de Stokes, con formas diferenciales, sus diferenciales exteriores, y la integración de éstas en cadenas. La "distancia" que hay entre estos dos hechos matemáticos es muy grande, y la intención del libro es proporcionar un plan de ruta al lector para que recorra el camino que separa estos dos hechos. En el transcurso de este principio y fin se exploran muchas de las maravillosas ideas que ofrece el cálculo en dimensiones superiores, como el concepto de diferenciabilidad de funciones reales de varias variables (capítulo 2), los teoremas de la función implícita y de la función inversa (capítulo 3), el problema de los extremos sujetos a restricciones (capítulo 4), los conceptos de curvatura y torsión para curvas en el espacio (capítulo 5), el teorema de cambio de variables en integrales dobles y triples (capítulo 6), el estudio de los campos conservativos y el teorema de Green (capítulo 7), los conceptos de superficies en el espacio (capítulo 8), el teorema de la divergencia y el teorema de Stokes (capítulo 9), y el teorema—general—de Stokes, como resultado globalizador de toda la obra (capítulo 10). Los temas mencionados, representativos de cada capítulo, constituyen un "guión" de un curso estándar de cálculo vectorial. Algunos de los temas adicionales que el libro presenta son: el teorema de Euler sobre funciones homogéneas (capítulo 2); el método de Newton para la solución de sistemas de ecuaciones no lineales (capítulo 3); un estudio sobre las condiciones que garantizan la existencia de extremos condicionados en el método de los multiplicadores de Lagrange (capítulo 4); un estudio de curvas paralelas (capítulo 5); el cálculo de volúmenes de esferas, conos y paralelepípedos en el espacio $\mathbb{R}^n$ (capítulo 6); un estudio introductorio sobre conjuntos conexos en $\mathbb{R}^n$, un estudio sobre las ecuaciones diferenciales exactas, y una demostración de la desigualdad isoperimétrica (capítulo 7); un estudio introductorio sobre tubos en $\mathbb{R}^2$ y $\mathbb{R}^3$ (capítulo 8); las "cuentas" explícitas para obtener la expresión del rotacional de un campo en el sistema de coordenadas esféricas (capítulo 9); la demostración del teorema general de Stokes, con formas diferenciales e integración en cadenas (capítulo 10); Además, un ejercicio con 27 incisos distribuídos en 4 secciones del libro (capítulo 2, secciones 6 y 12, y capítulo 7, secciones 3 y 4), en el que se dan algunas ideas sobre la teoría de funciones de variable compleja, y cuyo objetivo es que el lector aplique la teoría expuesta en esta obra para demostrar algunos resultados elementales que aparecen en esta teoría.

El libro contiene varios cientos de ejemplos resueltos y más de 2300 ejercicios para que el estudiante los resuelva, la mayoría de los cuales tiene respuesta en la sección correspondiente al

final del libro. El papel que juega la resolución de estos ejercicios en la comprensión del material expuesto es, como en todos los libros de matemáticas, fundamental. Hasta que nos enfrentamos a situaciones concretas planteadas en estos ejercicios, cuya solución demanda la aplicación de la teoría expuesta, es cuando se empieza a dar el proceso de comprensión de la materia. Los ejercicios que demandan para su solución algo más de lo que el libro ofrece, están marcados con uno o varios asteriscos, según su grado de dificultad.

Este libro fue escrito con el apoyo de una beca de Cátedra Patrimonial Nivel III del Consejo Nacional de Ciencia y Tecnología (CONACYT). Aunque la responsabilidad de la realización del proyecto fue solamente mía, en él estuvieron involucradas muchas personas que me ayudaron e impulsaron para presentar esta primera versión del libro, que inicialmente fue concebido como una obra menos ambiciosa de la que se presenta, pero que poco a poco se fue convirtiendo en lo que ahora es, al no poner resistencia a los encantos y ganas de escribir algunos de los temas complementarios del curso que se comentaban anteriormente. Antes que nada, deseo hacer patente mi agradecimiento a las autoridades de la Universidad Panamericana, que me ofrecieron el espacio y el apoyo para la realización de este proyecto; especialmente al Ing. Pedro Creuheras Vallcorba, de la Escuela de Ingeniería, y a la Lic. Aurea Rojas Ponce, del Centro de Cómputo, quienes siempre me brindaron las facilidades necesarias para salir adelante en los momentos críticos y decisivos del proyecto. Agradezco también al Girton College de la Universidad de Cambridge (Inglaterra), donde escribí los dos últimos capítulos del libro, durante el verano de 1994. A Sergio W. del Valle y Gutiérrez, quien trabajó conmigo durante medio año en una de las etapas finales del libro. A Carlos F. Diez de Sollano Navarro, a quien dirigí su tesis de licenciatura (sobre el producto cruz generalizado), algunos resultados de la cual aparecen en el ejercicio 35 de la sección 7 del capítulo 1. A Pedro Albin Smith, quien resolvió los ejercicios de los capítulos 5 y 6. Al Ing. Alfonso Leal Guajardo, quien revisó varios capítulos, usándolos en un curso sobre la materia que impartió en el primer semestre de 1994, y posteriormente revisó de manera exhaustiva el capítulo 7, resolviendo todos los ejercicios que en este capítulo aparecen. Al L.F.M. Francisco Ortíz Arango, al Ing. Eduardo de la Vega Segura, a la Ing. Lilia Elena de la Vega Segura, al Dr. Fernando Brambila Paz y al Dr. Alejandro Bravo Mojica, quienes leyeron varios de los capítulos del libro. Menciono de manera especial al equipo con quien trabajé durante las últimas horas antes de dar por concluido el proyecto, haciendo los dibujos del libro en computadora, armando, revisando, y, en fin, trabajando intensamente en esos momentos críticos de la terminación de un proyecto de esta magnitud; mi agradecimiento especial a mis alumnos Rigoberto Chávez Carrillo y José Luis Salazar Velázquez, al Ing. David Pérez Rivera, a la Ing. Lourdes Grimaldo Funes y a la Ing. Rebeca Moreno Lara Barragán. Por último, un agradecimiento más especial aún al Ing. Javier Cervantes Camarena, quien exhibió nuevamente una combinación muy difícil de conseguir, pues además de ayudarme con la elaboración de muchos de los dibujos que aparecen en el libro, logró, con su buen humor y optimismo, neutralizar muchos de mis momentos de histeria (que se incrementaron sustancialmente durante algunos meses previos a la terminación del libro), mostrándome siempre su amistad y apoyo.

Claudio de Jesús Pita Ruiz V.

Universidad Panamericana
Escuela de Ingeniería
Donatello 75-bis
Colonia Insurgentes-Mixcoac
México, D.F. 03920

México, D. F., septiembre de 1994.

Contenido

Capítulo 1

Introducción al espacio $\mathbb{R}^n$ y al álgebra lineal

En este primer capítulo expondremos los preliminares necesarios para abordar adecuadamente el estudio del cálculo para funciones cuyo dominio y/o codominio es el espacio n-dimensional $\mathbb{R}^n$. Por una parte, estudiaremos algunos aspectos sobre la naturaleza algebraica de este espacio, que será nuestro anfitrión durante el desarrollo de toda la obra, insistiendo en la gran riqueza geométrica, la cual puede ser visualizada en los casos en que $n = 2$ y $n = 3$ y, por otra parte, introduciremos algunos conceptos importantes del álgebra lineal que nos ayudarán en su momento a tener un lenguaje adecuado para entender varios de los temas que aparecen en el estudio del cálculo (sobre todo el diferencial) de las funciones anteriormente mencionadas (v.gr. la derivada de una función determinada es una "transformación lineal"). Advertimos, sin embargo, que los tópicos que aquí abordaremos no serán tratados en forma exhaustiva, pues el objetivo es solamente dejar asentado un material de repaso y/o referencia, cuyo conocimiento es importante (muchas veces fundamental) para entender las discusiones de los temas de esta obra. Muchos de estos temas se tratan de modo más profundo en algunos textos de álgebra lineal. De cualquier modo, se advierte que sí es un requisito el conocimiento de algunos resultados elementales sobre la teoría de sistemas de ecuaciones lineales, matrices y determinantes, que se exponen en los primeros capítulos de algunos libros de álgebra lineal, como por ejemplo, en los dos primeros capítulos de la referencia [PiI].

1.1 El espacio $\mathbb{R}^n$

Téngase en cuenta que, en todo el libro, la letra n, que acompaña a la letra $\mathbb{R}$ en la notación $\mathbb{R}^n$, denotará a un número natural.

Consideremos el conjunto de todas las n-adas ordenadas de números reales, que denotaremos por $\mathbb{R}^n$ (y leemos "erre ene")

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) | x_1, x_2, \ldots, x_n \in \mathbb{R}\}$$

A cada uno de los números reales $x_1, x_2, \ldots, x_n$ que conforman la n-ada $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, se le llama *componente* o *coordenada* de la n-ada correspondiente y, puesto que éstas son ordenadas, decimos, con más precisión, que x_i es la i-ésima coordenada de $(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$. Por ejemplo, si $n = 1$, el conjunto $\mathbb{R}^1$ no es más que el conjunto de números reales $\mathbb{R}$. Si $n = 2$, $\mathbb{R}^2$ será

el conjunto de parejas ordenadas de números reales que podemos escribir como $\{(x, y)|x, y \in \mathbb{R}\}$. Si $n = 3$, el conjunto $\mathbb{R}^3$ estará formado por las ternas ordenadas de números reales, que se puede escribir como $\{(x, y, z)|x, y, z \in \mathbb{R}\}$, etc. Insistimos en que las n-adas que constituyen el conjunto $\mathbb{R}^n$, son *ordenadas*: por ejemplo, en $\mathbb{R}^2$ la pareja $(2, 7)$ es *diferente* de la pareja $(7, 2)$. De hecho, dos n-adas de $\mathbb{R}^n$ se dicen ser *iguales*, cuando todas y cada una de sus coordenadas son iguales. Es decir que

$$(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n) \Leftrightarrow x_i = y_i, \qquad i = 1, 2, \ldots, n$$

Un hecho de fundamental importancia en el conjunto $\mathbb{R}^n$ es que podemos definir en él dos operaciones entre sus elementos, las cuales cumplen con ciertas propiedades que veremos a continuación. Este hecho hace que tal conjunto tenga una estructura algebraica llamada *espacio vectorial* y que, por tanto, nos podamos referir a él no sólo como el "conjunto $\mathbb{R}^n$", sino como el "espacio $\mathbb{R}^n$". Las operaciones que definimos en $\mathbb{R}^n$ son

a. Suma de n-adas ordenadas

Si $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$ son dos elementos de $\mathbb{R}^n$, definimos su *suma*, denotada por $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)$, como

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

b. Producto de una n-ada ordenada por un escalar

Si $(x_1, x_2, \ldots, x_n)$ es un elemento de $\mathbb{R}^n$, y c es un número real (en álgebra lineal se usa la palabra "escalar" para designar a un elemento de un campo, que en nuestro caso es $\mathbb{R}$; siguiendo esta nomenclatura, también nosotros llamaremos escalar a un número real). El producto de la n-ada $(x_1, x_2, \ldots, x_n)$ por el escalar c, denotada por $c(x_1, x_2, \ldots, x_n)$, se define como

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

Obsérvese que, según estas definiciones, tanto la suma de n-adas como el producto de una de ellas por un escalar, son nuevamente n-adas del conjunto $\mathbb{R}^n$. Por ello se dice que estas operaciones son *cerradas* en $\mathbb{R}^n$. Por ejemplo, en $\mathbb{R}^3$ la suma de la terna $(2, 5, -9)$ con la terna $(1, 0, 7)$ es $(2, 5, -9) + (1, 0, 7) = (2 + 1, 5 + 0, -9 + 7) = (3, 5, -2)$, que es una nueva terna de $\mathbb{R}^3$; en $\mathbb{R}^4$, el producto de la 4-ada $(2, 8, -6, 3)$ por el escalar -2 es $-2(2, 8, -6, 3) = (-4, -16, 12, -6)$ que es también un elemento de $\mathbb{R}^4$. Es fácil verificar que estas operaciones entre los elementos de $\mathbb{R}^n$ cumplen con las propiedades siguientes:

1. La suma es conmutativa, es decir

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (y_1, y_2, \ldots, y_n) + (x_1, x_2, \ldots, x_n)$$

2. La suma es asociativa, es decir

$$(x_1, x_2, \ldots, x_n) + \left[(y_1, y_2, \ldots, y_n) + (z_1, z_2, \ldots, z_n)\right] = \left[(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)\right] + (z_1, z_2, \ldots, z_n)$$

3. Existe un elemento en $\mathbb{R}^n$, llamado *cero*, que actúa de manera neutra para la suma. De hecho, este elemento es el que tiene todas sus coordenadas iguales al (número real) cero. Lo denotaremos

por $\mathbf{0}$ (o cuando no haya peligro de confusión, simplemente por 0). Es decir $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbb{R}^n$ y se tiene

$$(x_1, x_2, \ldots, x_n) + (0, 0, \ldots, 0) = (x_1, x_2, \ldots, x_n)$$

4. Cada n-ada de $\mathbb{R}^n$ tiene un "inverso aditivo", el cual es un elemento de $\mathbb{R}^n$ que tiene la propiedad de que, sumado con la n-ada original, produce cero (¡el cero de $\mathbb{R}^n$!). De hecho, el inverso aditivo de $(x_1, x_2, \ldots, x_n)$ es $(-x_1, -x_2, \ldots, -x_n)$ puesto que

$$(x_1, x_2, \ldots, x_n) + (-x_1, -x_2, \ldots, -x_n) = (0, 0, \ldots, 0)$$

5. Si λ es un escalar, se tiene

$$\lambda\left[(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)\right] = \lambda(x_1, x_2, \ldots, x_n) + \lambda(y_1, y_2, \ldots, y_n)$$

6. Si λ y μ son escalares, se tiene

$$(\lambda + \mu)(x_1, x_2, \ldots, x_n) = \lambda(x_1, x_2, \ldots, x_n) + \mu(x_1, x_2, \ldots, x_n)$$

7. Si λ y μ son escalares, se tiene

$$(\lambda\mu)(x_1, x_2, \ldots, x_n) = \lambda\left[\mu(x_1, x_2, \ldots, x_n)\right] = \mu\left[\lambda(x_1, x_2, \ldots, x_n)\right]$$

8. $1(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n)$

Como decíamos, todas estas propiedades son de verificación inmediata y su validez se basa fundamentalmente en las correspondientes propiedades ya conocidas de los números reales, como la conmutatividad, asociatividad y la existencia de neutros para la suma y producto de reales, además de la existencia de inverso para la suma y la distributividad. Obsérvese que, de hecho, si $n = 1$, $\mathbb{R}^1$ no es más que el conjunto de números reales y las propiedades 1–8 anteriormente mencionadas se cumplen automáticamente (basándonos en que conocemos de antemano las propiedades algebraicas de $\mathbb{R}$). Sabemos que, en realidad, $\mathbb{R}$ es más que un espacio vectorial: es un campo (es un hecho general que un campo K cualquiera es un espacio vectorial, tomando como escalares los mismos elementos de K).

A manera de ejemplo, verifiquemos la validez de la propiedad 5. Se tiene

$$\begin{aligned}
&\lambda\left[(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)\right] = \\
&\quad = \lambda(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) && \begin{pmatrix}\text{definición de}\\ \text{suma en } \mathbb{R}^n\end{pmatrix} \\
&\quad = \big(\lambda(x_1 + y_1), \lambda(x_2 + y_2), \ldots, \lambda(x_n + y_n)\big) && \begin{pmatrix}\text{definición de producto}\\ \text{por escalares en } \mathbb{R}^n\end{pmatrix} \\
&\quad = \big(\lambda x_1 + \lambda y_1, \lambda x_2 + \lambda y_2, \ldots, \lambda x_n + \lambda y_n\big) && \begin{pmatrix}\text{propiedad distributiva}\\ \text{de los números reales}\end{pmatrix} \\
&\quad = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) + (\lambda y_1, \lambda y_2, \ldots, \lambda y_n) && \begin{pmatrix}\text{definición de}\\ \text{suma en } \mathbb{R}^n\end{pmatrix} \\
&\quad = \lambda(x_1, x_2, \ldots, x_n) + \lambda(y_1, y_2, \ldots, y_n) && \begin{pmatrix}\text{definición de producto}\\ \text{por escalares en } \mathbb{R}^n\end{pmatrix}
\end{aligned}$$

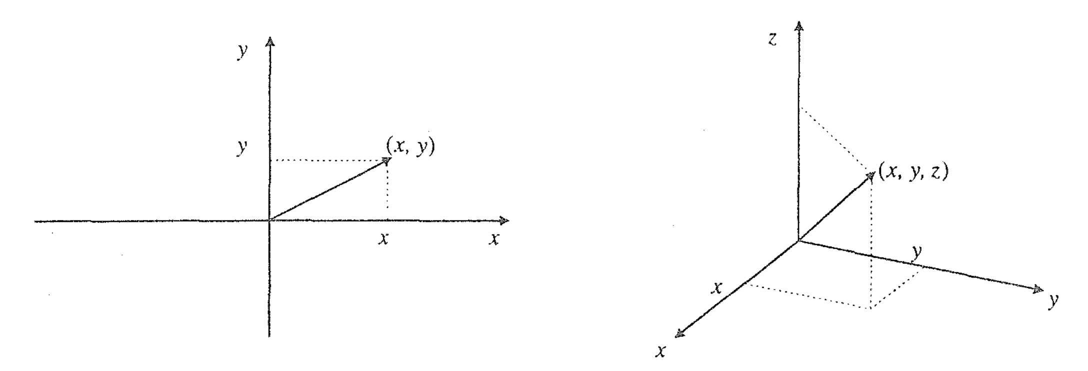


Figura 1. Vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$.

Cuando en un conjunto no vacío V se han definido operaciones de suma entre sus elementos y producto de éstos por escalares (números reales, o más en general, elementos de un campo K), y estas operaciones satisfacen (además de la cerradura) las propiedades 1–8 vistas anteriormente (es decir, la propiedad de conmutatividad de la suma, asociatividad de la suma, etc.), se dice que V es un espacio vectorial[1]. Así, el conjunto $\mathbb{R}^n$ se convierte en un espacio vectorial con las operaciones que en él hemos definido. De aquí en adelante nos referiremos a $\mathbb{R}^n$ como "el espacio $\mathbb{R}^n$" y a sus elementos (las n-adas ordenadas) como "vectores".

La *resta* de vectores en $\mathbb{R}^n$, digamos $\mathbf{x} - \mathbf{y}$, se define como

$$\mathbf{x} - \mathbf{y} = \mathbf{x} + (-\mathbf{y})$$

Cuando $n = 2$ ó $n = 3$, podemos visualizar geométricamente los espacios correspondientes $\mathbb{R}^2$ y $\mathbb{R}^3$. En efecto, dado un vector $\mathbf{v}$ en alguno de estos espacios, podemos ver a éste como el punto correspondiente del plano o del espacio tridimensional que tiene por coordenadas a las coordenadas de $\mathbf{v}$. Otro modo de verlo es como una flecha que parte del origen de coordenadas y llega al punto en cuestión. Más aún, toda "flecha" en el plano o en el espacio, puede ser pensada como un vector de $\mathbb{R}^2$ o $\mathbb{R}^3$, respectivamente. En efecto, supongamos que la flecha tiene su inicio en el punto $\mathbf{p}$ y su final en el punto $\mathbf{q}$. A ella asociamos entonces el vector $\mathbf{v} = \mathbf{q} - \mathbf{p}$ del espacio correspondiente. Con las

[1]Con más precisión, un espacio vectorial es un conjunto no vacío V en el cual están definidas dos operaciones entre sus elementos (llamados vectores), a saber, la suma de ellos $+: V \times V \to V$ con la cual a cada $\mathbf{v}_1, \mathbf{v}_2 \in V$ se le asocia un nuevo vector $(\mathbf{v}_1 + \mathbf{v}_2) \in V$, llamado "suma de $\mathbf{v}_1$ y $\mathbf{v}_2$", y el producto de un vector de V por un escalar (un elemento de un campo K, como $\mathbb{R}$ o $\mathbb{C}$) $\cdot: K \times V \to V$, con la cual, dado un $v \in V$ y un escalar $\lambda \in K$ (= $\mathbb{R}$ o $\mathbb{C}$), se le asocia un nuevo elemento $\lambda\mathbf{v} \in V$, llamado "producto del vector $\mathbf{v}$ por el escalar λ", cumpliendo las siguientes propiedades:
1. La suma es conmutativa: $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1, \forall \mathbf{v}_1, \mathbf{v}_2 \in V$.
2. La suma es asociativa: $\mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3) = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3, \forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$.
3. Existe en V un elemento neutro para la suma, llamado cero y denotado por $\mathbf{0}$. Es decir, existe $\mathbf{0} \in V$ tal que $\mathbf{v} + \mathbf{0} = \mathbf{v}$ $\forall \mathbf{v} \in V$.
4. Cada $\mathbf{v} \in V$ tiene asociado un inverso aditivo $(-\mathbf{v}) \in V$, con la propiedad de que $\mathbf{v} + (-\mathbf{v}) = 0$.
5. $\lambda(\mathbf{v}_1 + \mathbf{v}_2) = \lambda\mathbf{v}_1 + \lambda\mathbf{v}_2, \forall \lambda \in K, \forall \mathbf{v}_1, \mathbf{v}_2 \in V$.
6. $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}, \forall \lambda, \mu \in K, \forall \mathbf{v} \in V$
7. $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v}), \forall \lambda, \mu \in K, \forall \mathbf{v} \in V$.
8. $1\mathbf{v} = \mathbf{v}, \forall \mathbf{v} \in V$.
En este libro el espacio vectorial más importante con el que trabajaremos es justamente $\mathbb{R}^n$. Existen, sin embargo, otros espacios vectoriales importantes que eventualmente aparecerán en el desarrollo del libro, como el espacio de matrices, de funciones, etc.

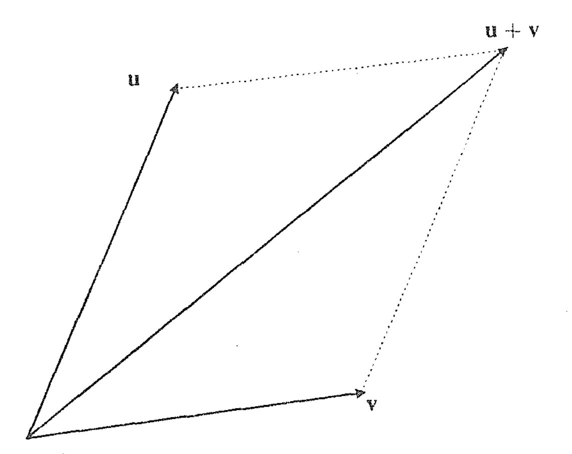


Figura 2. La suma de vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$.

consideraciones geométricas que veremos a continuación, será fácil ver que la flecha asociada a este vector $\mathbf{v}$, que parte del origen y llega al punto $\mathbf{q} - \mathbf{p}$, es "equivalente" (en el sentido de movimientos rígidos) a la flecha original que partía de $\mathbf{p}$ y llegaba a $\mathbf{q}$.

Debido a este tipo de identificaciones entre los puntos del plano cartesiano y del espacio tridimensional, con los vectores de los espacios vectoriales $\mathbb{R}^2$ y $\mathbb{R}^3$, es que se suele referir a estos espacios como "el plano $\mathbb{R}^2$" y "el espacio $\mathbb{R}^3$" respectivamente (refiriéndonos en este último caso al espacio tridimensional —en el que vivimos), y como ya lo decíamos en nuestro primer curso de cálculo "la recta $\mathbb{R}$".

Más aún, es interesante notar que las operaciones definidas en los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$ pueden ser visualizadas, al igual que algunas de las propiedades de ellas, con la ayuda de las versiones geométricas (las flechas) de los vectores de estos espacios. En efecto, se puede ver fácilmente (dejamos los detalles a cargo del lector) que la suma de vectores en estos espacios no es más que la "regla del paralelogramo" conocida en el manejo de flechas ("vectores geométricos") como se muestra en la figura 2.

Con ayuda de esta figura queda clara la validez de la propiedad conmutativa de la suma de vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$. También, usando esta idea, es fácil ver que la operación de resta de vectores, digamos $\mathbf{x} - \mathbf{y}$, equivale a tomar el vector (la flecha) que comienza en el punto $\mathbf{y}$ y termina en el punto $\mathbf{x}$ (el cual es en realidad una flecha que se obtiene por un mivimiento rígido de la flecha asociada a $\mathbf{x} - \mathbf{y}$).

Análogamente, con ayuda de la figura 4, queda clara la propiedad asociativa de la suma.

Por otra parte, la operación de producto por escalares puede también verse geométricamente de la siguiente manera: la multiplicación del vector $\mathbf{v}$ por el escalar λ produce un nuevo vector $\lambda\mathbf{v}$ (del que diremos que es un "múltiplo escalar" de $\mathbf{v}$) que, conservando la línea de acción de $\mathbf{v}$, se alarga (si $\lambda > 1$) o se contrae (si $0 < \lambda < 1$) manteniendo la misma dirección de $\mathbf{v}$, o invirtiendo tal dirección (si $\lambda < 0$). En particular, dado el vector $\mathbf{v} \in \mathbb{R}^2$ o $\mathbb{R}^3$, su inverso aditivo $(-\mathbf{v}) \in \mathbb{R}^2$ o $\mathbb{R}^3$ es una reproducción del vector $\mathbf{v}$ apuntando en la dirección "opuesta respecto del origen". Estos hechos se ilustran en la figura 5.

Todas las visualizaciones geométricas anteriores, a pesar de que sólo tienen sentido con vectores "que podemos ver", en los espacios $\mathbb{R}^2$ y/o $\mathbb{R}^3$, se acostumbra hacer uso de ellas en el caso general de vectores en $\mathbb{R}^n$, pensando en que de no tener las "limitaciones espaciales" que tenemos los seres

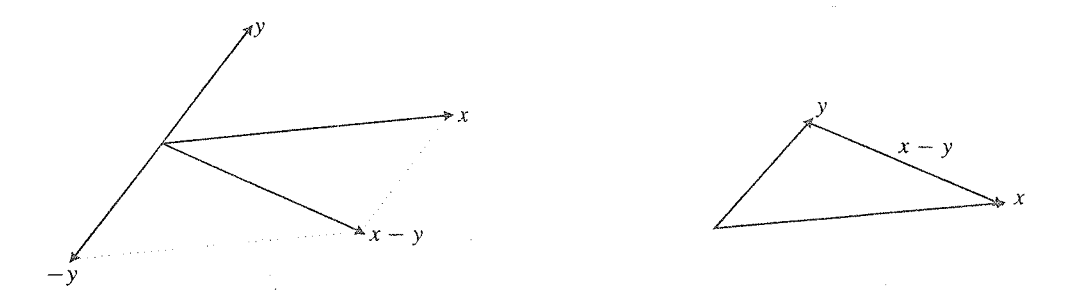


Figura 3. La resta de vectores $\mathbf{x} - \mathbf{y}$.

humanos (¡somos seres que vivimos en $\mathbb{R}^3$ y no podemos ver o imaginar espacios $\mathbb{R}^n$ con $n \geq 4$!), veríamos los vectores en $\mathbb{R}^n$ "con las mismas propiedades geométricas" que tienen los vectores en $\mathbb{R}^2$ o $\mathbb{R}^3$.

Algunas veces es importante considerar "pedazos" del espacio $\mathbb{R}^n$ que se comportan "algebraicamente de la misma manera" que el espacio total al que pertenecen. De hecho subconjuntos $S \subseteq \mathbb{R}^n$ que son en sí mismos espacios vectoriales con las operaciones de suma y producto por escalares que ya estaban definidas en $\mathbb{R}^n$ (es decir, que en S se cumplen la cerradura de las operaciones definidas en el espacio y las 8 propiedades que caracterizan a un espacio vectorial). Por ejemplo, si consideramos el subconjunto S de $\mathbb{R}^2$ dado por

$$S = \{(x, y) \in \mathbb{R}^2 | x = y\}$$

podemos verificar que los vectores de S satisfacen las 8 propiedades que cumple el espacio completo $\mathbb{R}^2$ que los hacen ser espacio vectorial: la cerradura (en S) de las operaciones de suma y producto por escalares se verifica fácilmente; que la suma es conmutativa y asociativa es un hecho que se cumple para *todos* los vectores de $\mathbb{R}^2$ y entonces, se cumple en particular para los vectores de S.

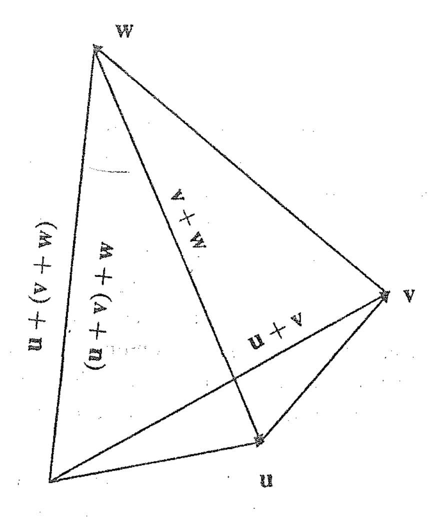


Figura 4. Versión geométrica de la propiedad asociativa de la suma de vectores.

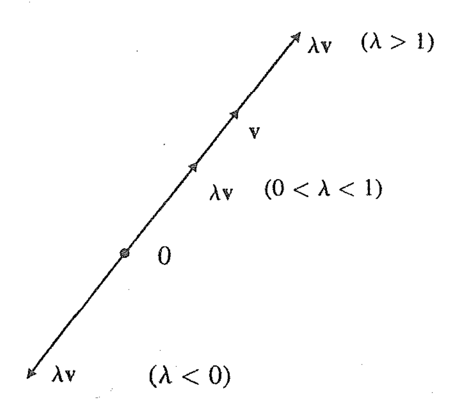


Figura 5. El producto del vector **v** por el escalar λ.

Que el vector **0** de $\mathbb{R}^2$ se encuentra en S es claro, pues tal vector es $(0, 0)$ (tiene sus dos coordenadas iguales). Dejamos que el lector verifique todas las demás propiedades. A un conjunto como éste se le llama subespacio de $\mathbb{R}^2$.

Generalmente, un *subespacio* de $\mathbb{R}^n$ es un subconjunto S de $\mathbb{R}^n$ que por sí mismo es un espacio vectorial. El siguiente resultado, que damos sin demostración, nos dice que es muy fácil saber cuando un subconjunto de $\mathbb{R}^n$ es un subespacio.

Teorema 1.1.1 Sea S un subconjunto no vacío del espacio $\mathbb{R}^n$. Entonces S es un subespacio de $\mathbb{R}^n$ si y solamente si: (1) dados $\mathbf{x}, \mathbf{y} \in S$, se tiene $\mathbf{x} + \mathbf{y} \in S$; (2) dado $\mathbf{x} \in S$, $c \in \mathbb{R}$, se tiene $c\mathbf{x} \in S$.

Demostración. Ejercicio. ∎

Ejemplo 1. Si consideramos el conjunto S que contiene solamente al vector **0** de $\mathbb{R}^n$, se tiene de inmediato que S es un subespacio de $\mathbb{R}^n$. También, si S es todo $\mathbb{R}^n$, es claro que será un subespacio. A estos dos subespacios de $\mathbb{R}^n$ se les llama *subespacios triviales.* El conjunto S de $\mathbb{R}^n$ dado por

$$S = \{(x_1, x_2, \ldots, x_n) | a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$$

donde $a_1, a_2, \ldots, a_n$ son números dados, es un subespacio de $\mathbb{R}^n$. En efecto, si $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ son vectores de S, se tiene que su suma $\mathbf{x} + \mathbf{y}$ también está en S, pues

$$\begin{aligned} a_1(x_1 + y_1) + a_2(x_2 + y_2) + \cdots + a_n(x_n + y_n) &= (a_1x_1 + a_2x_2 + \cdots + a_nx_n) \\ &+ (a_1y_1 + a_2y_2 + \cdots + a_ny_n) = 0 + 0 = 0 \end{aligned}$$

También, si $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ está en S y $c \in \mathbb{R}$, se tiene que $c\mathbf{x}$ está en S pues

$$a_1(cx_1) + a_2(cx_2) + \cdots + a_n(cx_n) = c(a_1x_1 + a_2x_2 + \cdots + a_nx_n) = c(0) = 0$$

En la sección 6 se verá que estos subespacios se pueden interpretar geométricamente como "hiperplanos" en $\mathbb{R}^n$. Otros subespacios importantes están dados por

$$S = \{(x_1, x_2, \ldots, x_n) | x_1 = a_1t, x_2 = a_2t, \ldots, x_n = a_nt\}$$

donde $a_1, a_2, \ldots, a_n$ son números reales dados y t es un real arbitrario. Dejamos que el lector verifique, usando el teorema anterior, que, efectivamente se trata de un subespacio de $\mathbb{R}^n$. Geométricamente estos subespacios se pueden identificar como "rectas que pasan por el origen" (como se verá en la sección 6). ∎

Dado un conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, decimos que el vector $\mathbf{v} \in \mathbb{R}^n$ es una *combinación lineal* de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, si existen escalares $c_1, c_2, \ldots, c_n$ tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Por ejemplo, el vector $(7, 5)$ es una combinación lineal de los vectores $(2, 1)$ y $(1, 1)$ puesto que $(7, 5) = 2(2, 1) + 3(1, 1)$; es decir, podemos escribir al vector $(7,5)$ como la suma de algún múltiplo escalar del vector $(2, 1)$ y algún múltiplo escalar del vector $(1, 1)$. Esto puede verse geométricamente como

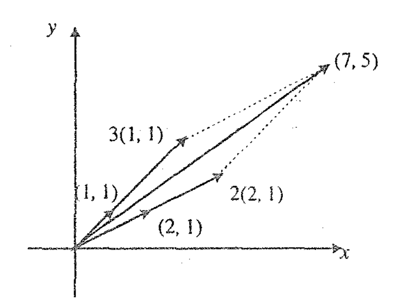


Figura 6. El vector $(7, 5) = 2(2, 1) + 3(1, 1)$.

En realidad, *cualquier* vector $(x, y) \in \mathbb{R}^2$ es una combinación lineal de los vectores $(2, 1)$ y $(1, 1)$. En efecto, podemos escribir $(x, y) = c_1(2, 1) + c_2(1, 1)$ con $c_1 = x - y$, $c_2 = 2y - x$, como se verifica sin dificultad.

Por otra parte, el vector $(1, 1, 0)$ *no* es una combinación lineal de los vectores $(1, 2, 3)$, $(-1, -1, 2)$, $(1, 3, 8)$. Para ver esto último, escribamos

$$(1, 1, 0) = c_1(1, 2, 3) + c_2(-1, -1, 2) + c_3(1, 3, 8)$$

y veamos que tales escalares c_1, c_2 y c_3 no existen. Haciendo las operaciones indicadas en la última expresión nos queda que

$$(1, 1, 0) = (c_1 - c_2 + c_3, 2c_1 - c_2 + 3c_3, 3c_1 + 2c_2 + 8c_3)$$

de donde se obtiene el sistema

$$c_1 - c_2 + c_3 = 1, \quad 2c_1 - c_2 + 3c_3 = 1, \quad 3c_1 + 2c_2 + 8c_3 = 0$$

del cual es fácil convencerse que no tiene solución.

Es claro que los vectores (2, 1) y (1, 1) tienen una propiedad importante que no tienen los vectores (1, 2, 3), (−1, −1, 2), (1, 3, 8), ya que con una combinación lineal adecuada de los primeros podemos escribir cualquier vector del espacio $\mathbb{R}^2$, cosa que no se puede hacer con los segundos vectores en el espacio $\mathbb{R}^3$. Tal propiedad es conocida como "independencia lineal" y a continuación haremos un estudio breve de ella, empezando por establecer la definición correspondiente.

Un conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ se dice ser *linealmente independiente* (abreviaremos l.i.) si la combinación lineal

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

obliga a que todos los escalares $c_1, c_2, \ldots, c_k$ sean cero. Es decir, si se tiene la implicación

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \Rightarrow c_1 = c_2 = \ldots = c_k = 0.$$

Caso contrario, se dice que los vectores son *linealmente dependientes* (abreviaremos l.d.)[2] Es decir, si se puede tener la combinación lineal $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ con no todos los escalares c_1, $c_2, \ldots, c_n$ iguales a cero.

Usando esta definición, es fácil convencerse de los siguientes hechos:

1. Cualquier conjunto de vectores que contenga al $\mathbf{0}$ es l.d.
2. Un conjunto formado sólo por un vector no nulo es l.i.
3. Si S es un conjunto de vectores l.i., cualquier subconjunto de S es también l.i.
4. Si S es un conjunto de vectores l.d., cualquier conjunto S' que contenga a S como subconjunto será también l.d.
5. Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son vectores l.d., entonces alguno de ellos se puede escribir como combinación lineal de los restantes.
6. Si $k > n$, el conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ es l.d.
7. Un conjunto de n vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ es l.i. si y sólo si el determinante de la matriz que tiene por vectores columna (o por vectores línea) a estos vectores es distinto de cero.

Dejamos al lector la verificación detallada de estos hechos (algunas de ellas usan resultados relacionados con los sistemas de ecuaciones lineales).

Un concepto muy importante que aparece cuando se trabaja en el espacio $\mathbb{R}^n$ es el concepto de base de este espacio. Se dice que un conjunto formado por n vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ es una *base* de $\mathbb{R}^n$, si estos vectores son linealmente independientes. Según la propiedad (7) anterior, los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ son (o forman) una base de $\mathbb{R}^n$ si y solamente si el determinante de la matriz cuyos vectores columna son los vectores dados, es distinto de cero. Esquemáticamente, los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ son una base de este espacio si y sólo si

$$\det \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \neq 0$$

[2] La propiedad de dependencia o independencia lineal se puede ver como una propiedad de los vectores o del conjunto que forman. No haremos distinción al respecto.

Cuando se tiene una base del espacio $\mathbb{R}^n$, digamos formada por el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, es importante considerar a éste último como un *conjunto ordenado* de vectores en $\mathbb{R}^n$. De esta manera, se tiene el siguiente resultado fundamental que pone de relieve la importancia de tener bases en el espacio $\mathbb{R}^n$.

Teorema 1.1.1 Si $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base del espacio $\mathbb{R}^n$, entonces cada vector $\mathbf{v} \in \mathbb{R}^n$ se escribe de manera única como combinación lineal de los vectores de β, es decir, existen únicos escalares $c_1, c_2, \ldots, c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$.

Demostración. Considere el conjunto $A = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$. Este es un conjunto linealmente dependiente, pues está formado por $n + 1 > n$ vectores de $\mathbb{R}^n$. Es decir, existen escalares $\gamma_1, \gamma_2, \ldots, \gamma_{n+1}$ no todos nulos tales que $\gamma_1\mathbf{v}_1 + \gamma_2\mathbf{v}_2 + \cdots + \gamma_n\mathbf{v}_n + \gamma_{n+1}\mathbf{v} = \mathbf{0}$. Afirmamos que $\gamma_{n+1} \neq 0$, pues caso contrario tendríamos $\gamma_1\mathbf{v}_1 + \gamma_2\mathbf{v}_2 + \cdots + \gamma_n\mathbf{v}_n = \mathbf{0}$ y, por la independencia lineal de los vectores de β, se concluiría que $\gamma_i = 0$ para $i = 1, 2, \ldots, n$, lo cual contradice la dependencia lineal del conjunto A. Tenemos entonces que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, donde $c_i = -\gamma_i/\gamma_{n+1}$. Veamos por último que estos escalares son únicos. Si existieran otros escalares tales que $\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n$, se tendría

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n$$

de donde

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \cdots + (c_n - d_n)\mathbf{v}_n = \mathbf{0}.$$

Usando la independencia lineal de la base β, concluímos de esta última expresión que $c_i - d_i = 0$, o sea que $c_i = d_i$ para todo $i = 1, 2, \ldots, n$, como se quería. Q.E.D.

En el teorema anterior, decimos que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ es "la expresión del vector $\mathbf{v}$ en términos de la base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$".

Ejemplo 2. Los vectores $\mathbf{v}_1 = (2, 1)$ y $\mathbf{v}_2 = (1, 1)$ (ver figura 6) forman una base de $\mathbb{R}^2$, puesto que ellos son l.i., hecho que se deduce del valor no nulo de $\det \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = 1$. En realidad, ya se había visto que todo vector $(x, y) \in \mathbb{R}^2$ se escribe (¡de manera única!) como $(x, y) = (x - y)\mathbf{v}_1 + (2y - x)\mathbf{v}_2$ como lo asegura el teorema anterior. ■

Ejemplo 3. Cualquier conjunto de k vectores en $\mathbb{R}^n$, con $k \neq n$, no puede ser una base de este espacio (¿por qué?). Los vectores $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (-1, -1, 2)$, $\mathbf{v}_3 = (1, 3, 8)$ no forman una base de $\mathbb{R}^3$ porque son l.d., ya que

$$\det \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 3 \\ 3 & 2 & 8 \end{bmatrix} = 0$$

(Obsérvese que este determinante ya había aparecido en una discusión previa sobre si *todo* vector de $\mathbb{R}^3$ se puede escribir como combinación lineal de los vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Se descubrió que no. Esto es justamente lo que volvimos a hacer en este ejercicio, ¿por qué?). ■

El ejemplo más importante de base en el espacio $\mathbb{R}^n$ es el conjunto $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ donde

$$\mathbf{e}_i = (0, \dots, 0, \underset{i-\text{ésima coordenada}}{\underset{\downarrow}{1}}, 0, \dots, 0) \qquad i = 1, 2, \dots, n.$$

Es claro que estos n vectores $\mathbf{e}_i$ constituyen una base de $\mathbb{R}^n$, pues la matriz cuyos vectores columna son estos vectores es justamente la matriz identidad (que tiene unos en su diagonal principal y ceros en las posiciones restantes), cuyo determinante es 1. Esta base es llamada *base canónica de* $\mathbb{R}^n$. Observe que el vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ se escribe en términos de esta base como

$$\mathbf{x} = (x_1, x_2, \dots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n.$$

Ejemplo 4. En el espacio $\mathbb{R}^2$, los vectores de la base canónica son $\mathbf{e}_1 = (1, 0)$ y $\mathbf{e}_2 = (0, 1)$, comúnmente denotados por $\mathbf{i}$ y $\mathbf{j}$, respectivamente. Entonces, dado cualquier vector $(x, y) \in \mathbb{R}^2$, éste se escribe en términos de esta base como

$$(x, y) = x\mathbf{i} + y\mathbf{j}.$$

Análogamente, los vectores de la base canónica del espacio $\mathbb{R}^3$, denotados por $\mathbf{i}, \mathbf{j}, \mathbf{k}$, son $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$, de modo que para $(x, y, z) \in \mathbb{R}^3$ se tiene

$$(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Geométricamente esto se ve de la siguiente manera

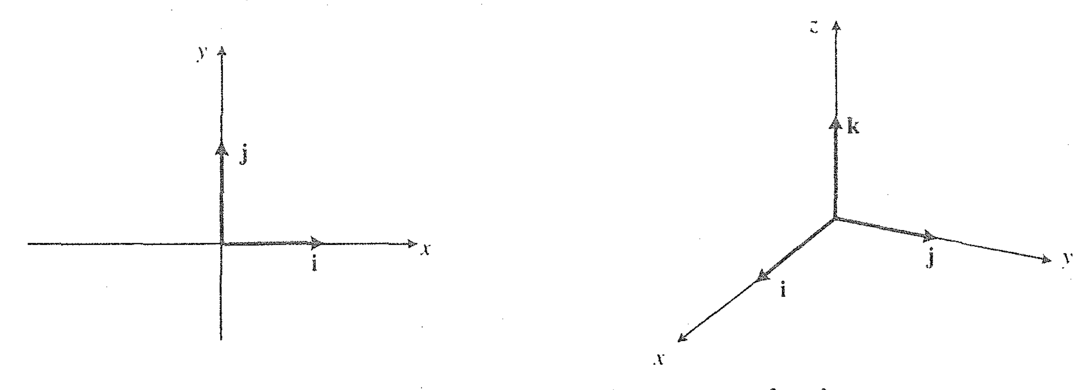


Figura 7. Los vectores de las bases canonicas de $\mathbb{R}^2$ y $\mathbb{R}^3$.

Cuando se consideran subespacios S de $\mathbb{R}^n$ es importante tener también un concepto de base de ellos, pues esto nos permitirá introducir el concepto de "dimensión del subespacio". Según se ha visto anteriormente, son dos las características de conjunto β que es base del espacio $\mathbb{R}^n$: (1) sus vectores son l.i. y (2) todo vector de $\mathbb{R}^n$ se puede escribir como combinación lineal de los vectores de β. Estas son, de hecho, las dos propiedades que *definen* una base de cualquier espacio vectorial.

Diremos entonces que un conjunto $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ del subespacio S de $\mathbb{R}^n$ es una base de S, si: (1) los vectores de β son l.i.; (2) todo vector $\mathbf{v} \in S$ se escribe como combinación lineal de los vectores de β. En este caso se dice que los vectores de β *generan* al subespacio S, o bien, que S es *generado* por los vectores de β.

Ejemplo 5. Sea $S = \{(x, y)|x = y\}$. Ya se vió que S es un subespacio de $\mathbb{R}^2$. Es fácil verificar que el conjunto $\beta = \{(1, 1)\}$ es una base de S, pues, por una parte es claro que es l.i. y, por otra, todo vector de S, digamos (a, a) es un múltiplo escalar de $(1, 1)$. ■

Ejemplo 6. El conjunto $S = \{(x, y, z)|x - 3y - 8z = 0\}$ es un subespacio de $\mathbb{R}^3$ (ver ejemplo 1). Para "descubrir" una base de S podemos proceder como sigue: escribamos un vector arbitrario de S y tratemos de "descomponerlo" como combinación lineal de vectores de $\mathbb{R}^3$. Verificando finalmente que estos vectores son l.i. podemos concluir que constituyen la base buscada. En nuestro caso tenemos que un vector cualquiera de S se escribe como

$$(x, y, z) = (3y + 8z, y, z) = y(3, 1, 0) + z(8, 0, 1)$$

Entonces, todo vector de S se escribe como combinación lineal de los vectores $\mathbf{v}_1 = (3, 1, 0)$ y $\mathbf{v}_2 = (8, 0, 1)$. Puesto que estos vectores son l.i., concluimos que $\beta = \{(3, 1, 0), (8, 0, 1)\}$ es una base de S. ■

Por supuesto que la base β de un subespacio S de $\mathbb{R}^n$ no es única. Con el mismo ejemplo anterior, podemos ver que $\beta_2 = \{(-2, 2, -1), (-1, 3, -1)\}$ es otra base del subespacio $S = \{(x, y, z)|x - 3y - 8z = 0\}$. Lo que sí ocurre, y es posible demostrar en general, es que dos bases distintas de un subespacio *tienen siempre* el *mismo número de vectores*. Esto nos permite establecer la siguiente definición importante.

> **Definición.** Sea S un subespacio de $\mathbb{R}^n$. Se llama *dimensión* de S, denotada por dim S, al número de vectores que existe en una base (cualquiera) de S. ■

Por ejemplo, es claro que la dimensión de $\mathbb{R}^n$ es n. La dimensión del subespacio del ejemplo 6 es 2. En el caso del subespacio trivial $\{\mathbf{0}\}$ que contiene solamente al vector cero, su dimensión se define como siendo cero.

Ejercicios (Capítulo 1, Sección 1)

1. Verifique que el conjunto de n-adas ordenadas de números reales

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n)|x_i \in \mathbb{R}, \qquad i = 1, 2, \dots, n\}$$

con las operaciones de suma y producto por escalares definidas en esta sección, es un espacio vectorial.

2. Escriba en forma explícita.
   - **a.** el neutro para la suma de $\mathbb{R}^3$.
   - **b.** el inverso aditivo de $(1, 2, -3, 5) \in \mathbb{R}^4$.
   - **c.** el inverso aditivo del inverso aditivo de un vector $\mathbf{v} \in \mathbb{R}^n$.
   - **d.** el inverso aditivo del neutro para la suma de un vector $\mathbf{v} \in \mathbb{R}^n$.
   - **e.** el vector suma de $(1, 1, 1)$ con $(3, 2, 2)$ en $\mathbb{R}^3$.
   - **f.** la propiedad conmutativa para la suma de vectores en $\mathbb{R}^3$.

g. el vector suma de (8, 9, 3, 5) con el inverso aditivo de (2, 7, 5, 4) en $\mathbb{R}^4$.
h. el vector suma de $(a, b, c, d, e) \in \mathbb{R}^5$ con su inverso aditivo.
i. 3 veces el vector (2, 1, 1) de $\mathbb{R}^3$.
j. -5 veces el vector $(1, 1, 1, 1, 1) \in \mathbb{R}^5$.
k. la propiedad asociativa para la suma de vectores en $\mathbb{R}^4$.
l. el inverso aditivo de 4 veces el vector $(2, 4, -7) \in \mathbb{R}^3$.
m. la suma de (2, 5, 5, 4) con 4 veces el vector $(-1, -2, -1, -1)$ en $\mathbb{R}^4$.
n. la suma de (1, 1) con el inverso aditivo de 5 veces el vector (4, 5) en $\mathbb{R}^2$.
o. la suma del inverso aditivo de (1, 1) con 5 veces el vector (4, 5) en $\mathbb{R}^2$.
p. el vector $3(1, 1, 8) + 4\left[(-2, 3, 0) + 5(1, 0, 1)\right]$ de $\mathbb{R}^3$.
q. el vector $-(2, 0, 2) + 3\{(3, 2, 2) - 3[-(1, 2, 1) + 7(0, 1, 0]\}$ de $\mathbb{R}^3$.
r. el vector $(2, 1, 0, 0) - 2(1, 1, 1, 1)$ de $\mathbb{R}^4$ multiplicado por el escalar -5.
s. el inverso aditivo del vector $-(1, 4, 2, 3) + 2(3, 2, 1, 1)$ en $\mathbb{R}^4$ multiplicado por el escalar -6.
t. el vector de $\mathbb{R}^4$ que sumado al vector (3, 2, 0, 0) dé por resultado el vector (1, 1, 2, 1).
u. el vector de $\mathbb{R}^3$ que sumado con el inverso aditivo del vector $(1, -4, 6)$ da por resultado el vector 3(3, 4, 2).

3. Sean $\mathbf{x}$ y $\mathbf{y}$ dos vectores de $\mathbb{R}^n$. Verifique que el inverso aditivo del vector $\mathbf{x} - \mathbf{y}$ es el vector $\mathbf{y} - \mathbf{x}$. Discuta geométricamente este hecho.

4. Demuestre que el subconjunto de $\mathbb{R}^2$

$$S = \{(x, y)|x = y\}$$

con las operaciones usuales de suma y producto por escalares (las que están definidas en $\mathbb{R}^2$) es un espacio vectorial, verificando que se cumplen los 8 axiomas que definen a esta estructura algebraica. Observe que geométricamente este conjunto es representado por la recta $y = x$, la cual es una recta que pasa por el origen.

5. Usando el teorema 1.1.1, demuestre que el conjunto de vectores en $\mathbb{R}^2$ que se encuentran en la recta que pasa por el origen $ax + by = 0$ (con a y b reales no ambos nulos), es un subespacio de $\mathbb{R}^2$. Más aún, demuestre que si S es un subespacio *no trivial* de $\mathbb{R}^2$ (es decir, distinto de $\{(0, 0)\}$ y de todo $\mathbb{R}^2$), entonces S es una recta que pasa por el origen, siguiendo los pasos:

a. Tome un vector $(x_0, y_0) \in S$ no nulo, digamos que $x_0 \neq 0$, y defina

$$L = \{(x, y)|y_0x - x_0y = 0\}.$$

Ciertamente L es un subespacio de $\mathbb{R}^2$ (¿por qué?). Tome $(x_1, y_1) \in L$. Verifique que (x_1, y_1) es un múltiplo escalar de (x_0, y_0). Con S como un subespacio de $\mathbb{R}^2$, concluya que (x_1, y_1) pertenece de hecho a S. Esto demuestra que $L \subset S$.

b. Suponga que S no está contenido en L. Tome entonces un vector $(x_2, y_2) \in S$ tal que $(x_2, y_2) \notin L$ (es decir, $y_0x_2 - x_0y_2 \neq 0$). Considere el sistema de dos ecuaciones con las incógnitas u y v,

$$y_0u + y_2v = y, \qquad x_0u + x_2v = x$$

donde x, y y son números reales arbitrarios dados. Concluya que este sistema de ecuaciones tiene solución única, digamos $\tilde{u}$, $\tilde{v}$. Verifique entonces que el vector $(x, y) \in \mathbb{R}^2$ se puede escribir como

$$(x, y) = \tilde{u}(x_0, y_0) + \tilde{v}(x_2, y_2).$$

Use S como un subespacio de $\mathbb{R}^2$ para concluir que $(x, y) \in S$, con lo cual concluya a su vez que $S = \mathbb{R}^2$, lo que es una contradicción a las hipótesis hechas sobre S. Esto prueba entonces que $S \subset L$.

c. Tome los resultados de los dos incisos anteriores para concluír finalmente que $S = L$, y que, entonces, S es una recta que pasa por el origen.

6. Demuestre que dos vectores en $\mathbb{R}^n$ son linealmente dependientes si y sólo si uno de ellos es un múltiplo del otro.

7. Use el resultado del ejercicio anterior para decidir (a simple vista) si los siguientes pares de vectores son linealmente independientes o dependientes.

a. $(1, 1)$ y $(2, 3)$.

b. $(2, 4, 1)$ y $(8, 16, 4)$.

c. $(0, 0, 0)$ y $(3, 2, -7)$.

d. $(1, 1, 2, 0, 1)$ y $(3, 3, 6, 0, 3)$.

e. $(2, 5, 1, 0, 1)$ y $(-3, 5, 1, 0, 1)$.

8. Demuestre que cualquier conjunto de vectores en $\mathbb{R}^n$ que contenga al vector cero es linealmente dependiente. (Sugerencia: use directamente la definición de dependencia lineal).

9. Pruebe que un conjunto formado por un solo vector no nulo es linealmente independiente.

10. Demuestre que si S es un conjunto de vectores en $\mathbb{R}^n$ linealmente indpendiente, entonces cualquier subconjunto de S es también linealmente independiente.

11. Demuestre que si S es un conjunto de vectores en $\mathbb{R}^n$ linealmente dependiente, entonces cualquier conjunto que contenga a S es también linealmente dependiente.

12. Pruebe que si los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ son linealmente dependientes, entonces alguno de ellos se puede escribir como combinación lineal de los restantes.

13. Demuestre que si $k > n$, el conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ es linealmente dependiente. (Sugerencia: escriba explícitamente la combinación lineal $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ con las coordenadas de los vectores involucrados; obtendrá un sistema homogéneo de n ecuaciones lineales con k indeterminadas $c_1, c_2, \ldots, c_k$. Use el hecho de que para un sistema de este tipo, si $k > n$, existen soluciones no todas nulas para las incógnitas).

14. Demuestre que un conjunto formado por n vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ es linealmente independiente si y solamente si la matriz cuadrada de orden n que tiene por vectores columna (o por vectores línea) a estos vectores, tiene determinante distinto de cero. (Sugerencia: escriba explícitamente la combinación lineal $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0$ con las coordenadas de los vectores involucrados; obtendrá así un sistema homogéneo de n ecuaciones lineales con n incógnitas $c_1, c_2, \ldots, c_n$. Use el hecho de que un sistema semejante tiene sólo la solución trivial $c_1 = c_2 = \cdots = c_n = 0$ si y sólo si el determinante de la matriz del sistema —que es el mismo que el de su transpuesta— es no nulo).

15. Diga si los siguientes conjuntos de vectores son linealmente independientes o dependientes, justificando su respuesta directamente de la definición, o bien, usando alguno de los resultados de los problemas 8–14 anteriores.

a. $\{(2, 1)\}$.

b. $\{(3, 2, 1), (1, 0, 0), (-4, 5, -2)\}$.

c. $\{(1, 1, 9), (2, 1, 3), (2, 2, 3), (3, -3, -7)\}$.

d. $\{(1, 4, 5, 0), (2, 1, 0, 0), (3, 1, 1, 1)\}$.

e. $\{(1, 2, 3, 4), (0, 2, 3, 4), (0, 0, 3, 4), (0, 0, 0, 4)\}$.

16. Demuestre el teorema 1.1.1. (Sugerencia: el "sólo si" es obvio; para probar el "si", observe que la cerradura de las operaciones en el espacio vectorial queda garantizada por las dos condiciones dadas; la conmutatividad y asociatividad de la suma, y las propiedades relacionadas con productos por escalares se cumplen automáticamente —¿por qué?—; resta por ver que existe el neutro para la suma en S y que cada x de S tiene en S su inverso aditivo; esto lo puede hacer usando la propiedad 2. con $c = 0$ y $c = -1$).

17. Diga si cada uno de los siguientes conjuntos son subespacios del espacio $\mathbb{R}^n$ correspondiente.

a. $S = \{(x, y)|2x + y = 0\} \subset \mathbb{R}^2$.

b. $S = \{(x, y, z)|2x + y = 0\} \subset \mathbb{R}^3$.

c. $S = \{(x, y, z)|x^2 + y = 0\} \subset \mathbb{R}^3$.

d. $S = \{(x, y, z)|x^2 + y^2 + z^2 = 0\} \subset \mathbb{R}^3$.

e. $S = \{(x, y, z)|x^2 + y^2 + z^2 \geq 0\} \subset \mathbb{R}^3$.

f. $S = \{(x, y, z)|x^2 + y^2 + z^2 > 0\} \subset \mathbb{R}^3$.

g. $S = \{(x_1, x_2, x_3, x_4)|x_1 = x_2 = x_3 = x_4\} \subset \mathbb{R}^4$.

h. $S = \{(x_1, x_2, x_3, x_4)|x_1x_2x_3x_4 = 1\} \subset \mathbb{R}^4$.

i. $S = \{(x_1, x_2, x_3, x_4, x_5)|x_1 + x_2 + x_3 + x_4 + x_5 = a\} \subset \mathbb{R}^5$ (a un número dado).

18. Para cada uno de los subconjuntos S de $\mathbb{R}^3$ dados a continuación verifique que se trata de subespacios y encuentre una base de ellos, así como su dimensión.

a. $S = \{(x, y, z)|z = 0\}$.

b. $S = \{(x, y, z)|x = y = 0\}$.

c. $S = \{(x, y, z)|x + y = 0\}$.

d. $S = \{(x, y, z)|x + y + z = 0\}$.

e. $S = \{(x, y, z)|3x - 8y + 9z = 0\}$.

f. $S = \{(x, y, z)|x = 2t,\ y = t,\ z = 5t,\ t \in \mathbb{R}\}$.

g. $S = \{(x, y, z)|x = 2y = 3z\}$.

h. $S = \{(x, y, z)|x = s,\ y = 8s,\ z = 0,\ s \in \mathbb{R}\}$.

i. $S = \{(x, y, z)|x = y\}$.

19. Explique por qué cada uno de los siguientes conjuntos de vectores de $\mathbb{R}^3$ no pueden constituír una base de este espacio.

a. $\{(1, 2, 1), (2, 5, 4)\}$.

b. $\{(1, -1, 3), (0, 0, 0), (2, 3, 6)\}$.

c. $\{(2, 5, 4), (1, 3, 2), (2, 6, 4)\}$.
d. $\{(1, 1, 5), (1, 1, 5), (3, 2, 2)\}$.
e. $\{(3, 2, 1), (2, 5, 5), (3, 4, 2), (2, 2, 7)\}$.
f. $\{(2, 4, 8)\}$.
g. $\{(3, 4, 3), (1, 1, 1), (2, 2, 2)\}$.
h. $\{(1, 1, 3), (2, 6, 4), (5, 3, 5), (3, 2, 1), (2, 3, 7)\}$.

20. Verifique que todo vector de $\mathbb{R}^2$ se puede escribir como una combinación lineal de los vectores $\mathbf{v}_1 = (1, 3)$, $\mathbf{v}_2 = (3, 7)$, $\mathbf{v}_3 = (-3, 5)$. ¿Significa esto que el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base de $\mathbb{R}^2$?

21. Verifique que los siguientes conjuntos constituyen bases de los espacios correspondientes.
a. $\{(1, 2), (3, 1)\}$ de $\mathbb{R}^2$.
b. $\{(1, 1), (9, 11)\}$ de $\mathbb{R}^2$.
c. $\{(1, 1, 1), (0, 5, 2), (0, 0, 19)\}$ de $\mathbb{R}^3$.
d. $\{(1, -1, -1), (2, 3, 1), (2, 7, 3)\}$ de $\mathbb{R}^3$.
e. $\{(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)\}$ de $\mathbb{R}^4$.
f. $\{(2, 3, 4, 2, 3), (0, 2, 4, 3, 5), (0, 0, 1, 0, 0), (0, 0, 0, 4, 2), (0, 0, 0, 0, 3)\}$ de $\mathbb{R}^5$.

22. Demuestre que el conjunto $\beta = \{(a, b), (c, d)\}$ es una base de $\mathbb{R}^2$ si y sólo si $ad - bc \neq 0$.

23. Verifique que $\beta = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ es una base del espacio $\mathbb{R}^3$. Escriba el vector (x, y, z) en términos de esta base.

24. Demuestre la afirmación recíproca del teorema 1.1.2. Es decir, demuestre que si el conjunto $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de n vectores en $\mathbb{R}^n$ es tal que cada vector $\mathbf{v} \in \mathbb{R}^n$ se escribe de manera única como combinación lineal de los vectores de β, entonces β es una base de $\mathbb{R}^n$. (Sugerencia: la expresión $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ es una manera de escribir el vector $\mathbf{0} \in \mathbb{R}^n$. Otra manera es la trivial $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n$. Obtenga de aquí la independencia lineal de β . . .). Concluya entonces que las dos afirmaciones siguientes acerca del conjunto β son equivalentes:
a. β es una base de $\mathbb{R}^n$ (es decir, β es un conjunto linealmente independiente).
b. todo vector $\mathbf{v} \in \mathbb{R}^n$ se escribe de manera única como combinación lineal de los vectores de β.

25. Considere los vectores $\mathbf{v}_1 = (x_1, y_1)$, $\mathbf{v}_2 = (x_2, y_2)$ en $\mathbb{R}^2$. Defina el *producto de* $\mathbf{v}_1$ *por* $\mathbf{v}_2$, denotado por $\mathbf{v}_1\mathbf{v}_2$, coordenada a coordenada (como se hizo con la suma). Es decir, defina $\mathbf{v}_1\mathbf{v}_2 = (x_1y_1, x_2y_2)$. Observe que $\mathbf{v}_1\mathbf{v}_2$ es un nuevo vector de $\mathbb{R}^2$. Demuestre que:
a. el producto es conmutativo. Es decir, $\mathbf{v}_1\mathbf{v}_2 = \mathbf{v}_2\mathbf{v}_1$, $\forall \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$.
b. el producto es asociativo. Es decir, $\mathbf{v}_1(\mathbf{v}_2\mathbf{v}_3) = (\mathbf{v}_1\mathbf{v}_2)\mathbf{v}_3$, $\forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^2$.
c. existe un neutro para el producto $\mathbf{1} \in \mathbb{R}^2$, tal que $\mathbf{v}\mathbf{1} = \mathbf{v}$, $\forall \mathbf{v} \in \mathbb{R}^2$.
d. el producto es distributivo. Es decir, $\mathbf{v}_1(\mathbf{v}_2 + \mathbf{v}_3) = \mathbf{v}_1\mathbf{v}_2 + \mathbf{v}_1\mathbf{v}_3$, $\forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^2$ (con la suma definida en esta sección).
e. Si $\mathbf{v} = (a, b)$ es un vector cualquiera de $\mathbb{R}^2$, entonces $\mathbf{v} = \mathbf{v}\mathbf{i} + \mathbf{v}\mathbf{j}$, en donde $\mathbf{i} = (1, 0)$, $\mathbf{j} = (0, 1)$.

f. ¿Existe un vector inverso multiplicativo asociado a todo vector no nulo (es decir, distinto del vector (0, 0)) $\mathbf{v} \in \mathbb{R}^2$?. Es decir, dado $\mathbf{v} \in \mathbb{R}^2$ no nulo, ¿existe $\mathbf{v}^{-1} \in \mathbb{R}^2$ tal que $\mathbf{v}^{-1}\mathbf{v} = \mathbf{1}$ (el vector neutro multiplicativo de $\mathbb{R}^2$ del inciso c))?

g. ¿Vale la ley de la cancelación para este producto definido en $\mathbb{R}^2$?. Es decir, ¿es cierto que si $\mathbf{v}_1\mathbf{v}_2 = \mathbf{v}_1\mathbf{v}_3$ y $\mathbf{v}_1$ es distinto de (el vector) cero, entonces $\mathbf{v}_2 = \mathbf{v}_3$?

1.2 Producto punto. Proyecciones

En el espacio $\mathbb{R}^n$ podemos definir un tipo de producto entre sus elementos (los vectores del espacio) con el cual este espacio se llena de una gran riqueza geométrica que nos permite adentrarnos más en la "esencia" misma de la naturaleza de él. Este producto es el conocido "producto punto", el cual no es más que un tipo de "producto interno" que se puede definir en un espacio vectorial en general.

El *producto punto* en $\mathbb{R}^n$ es una función $\cdot: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ que a cada par de vectores $\mathbf{x}$ y $\in \mathbb{R}^n$ le asocia un número real $\mathbf{x} \cdot \mathbf{y}$ (llamado también "producto punto de $\mathbf{x}, \mathbf{y}$"; se usa también la notación $\langle \mathbf{x}, \mathbf{y} \rangle$) dado por

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

en el que $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$.

En el teorema siguiente se recogen las propiedades más importantes del producto punto.

Teorema 1.2.1 El producto punto $\mathbf{x} \cdot \mathbf{y}$ de dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ tiene las siguientes propiedades:

1. $\mathbf{x} \cdot \mathbf{x} \geq 0$, $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
2. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
3. $(\mathbf{x} + \mathbf{x}') \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} + \mathbf{x}' \cdot \mathbf{y}$
4. $(c\mathbf{x}) \cdot \mathbf{y} = c(\mathbf{x} \cdot \mathbf{y})$, $c \in \mathbb{R}$

Demostración. Se trata de verificaciones de simple rutina. Hacemos las cuentas correspondientes a las propiedades (3) y (4) (simultáneamente) y dejamos que el lector haga las de las propiedades (1) y (2). Si $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{x}' = (x'_1, x'_2, \ldots, x'_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ son tres vectores cualesquiera de $\mathbb{R}^n$ y $c \in \mathbb{R}$, se tiene

$$\begin{aligned}(c\mathbf{x} + \mathbf{x}') \cdot \mathbf{y} &= (cx_1 + x'_1)y_1 + (cx_2 + x'_2)y_2 + \cdots + (cx_n + x'_n)y_n \\ &= c(x_1y_1 + x_2y_2 + \cdots + x_ny_n) + (x'_1y_1 + x'_2y_2 + \cdots + x'_ny_n) \\ &= c(\mathbf{x} \cdot \mathbf{y}) + \mathbf{x}' \cdot \mathbf{y}\end{aligned}$$

Q.E.D

El teorema anterior nos dice que el producto punto es una función *definida positiva* (propiedad 1), *simétrica* (propiedad 2), y lineal respecto de su primera variable (propiedades 3 y 4). Juntando este último hecho, con la simetría del producto punto, concluimos que éste es lineal también respecto de su segunda variable, de modo que es entonces una función *bilineal*. Esta propiedad de bilinealidad se usa frecuentemente en la forma

$$\left(\sum_i c_i\mathbf{u}_i\right) \cdot \left(\sum_j d_j\mathbf{v}_j\right) = \sum_i \sum_j c_i d_j (\mathbf{u}_i \cdot \mathbf{v}_j)$$

donde $\mathbf{u}_i$, $\mathbf{v}_j$ son vectores de $\mathbb{R}^n$ y c_i, d_j son escalares.

En el siguiente teorema se establece una de las desigualdades más célebres de la matemática, en su versión para vectores en el espacio $\mathbb{R}^n$.

Teorema 1.2.2 **(Desigualdad de Cauchy-Schwarz)** Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores cualesquiera en $\mathbb{R}^n$. Entonces

$$(\mathbf{x}\cdot\mathbf{y})^2 \leq (\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y})$$

Demostración. Si $\mathbf{x} = \mathbf{0}$, ambos miembros de la desigualdad son iguales a cero (y entonces en este caso es cierta la desigualdad). Sea entonces $\mathbf{x} \neq \mathbf{0}$. Considere el vector $\mathbf{u} = \mathbf{y} + c\mathbf{x}$, donde c es un número real fijo, pero arbitrario. Consideremos el producto punto de $\mathbf{u}$ con $\mathbf{u}$ y apliquemos las propiedades del teorema 1.2.1, para obtener que

$$0 \leq (\mathbf{y} + c\mathbf{x})\cdot(\mathbf{y} + c\mathbf{x}) = (\mathbf{y}\cdot\mathbf{y}) + 2(\mathbf{x}\cdot\mathbf{y})c + (\mathbf{x}\cdot\mathbf{x})c^2$$

La función $\psi(c) = (\mathbf{x}\cdot\mathbf{x})c^2 + 2(\mathbf{x}\cdot\mathbf{y})c + (\mathbf{y}\cdot\mathbf{y})$ es una función cuadrática que geométricamente representa una parábola que abre hacia arriba (ya que $\mathbf{x}\cdot\mathbf{x}$ es positivo), y, puesto que $\psi(c) \geq 0$ para toda c, no debe *cortar* al eje de las c. Es decir, el discriminante de la ecuación cuadrática

$$(\mathbf{y}\cdot\mathbf{y}) + 2(\mathbf{x}\cdot\mathbf{y})c + (\mathbf{x}\cdot\mathbf{x})c^2 = 0$$

debe ser no positivo (i.e. debe ser o negativo —caso en el que la parábola no toca al eje de abscisas—, o cero —caso en el que la parábola *toca* al eje de abscisas). Entonces

$$(2(\mathbf{x}\cdot\mathbf{y}))^2 - 4(\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y}) \leq 0$$

o sea

$$(\mathbf{x}\cdot\mathbf{y})^2 \leq (\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y})$$

como se quería demostrar. Q.E.D.

Haciendo explícitos los productos punto involucrados en la desigualdad de Cauchy-Schwarz, ésta se convierte en una desigualdad de números: si $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ son números arbitrarios, entonces

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

Para motivar la definición que daremos de "ortogonalidad" de vectores en $\mathbb{R}^n$, estudiemos cómo aparece el hecho de que dos vectores en $\mathbb{R}^2$ sean perpendiculares. Consideremos los vectores $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$ en $\mathbb{R}^2$ y supongamos que son perpendiculares, como se muestra en la siguiente figura
La recta en la que se encuentra el vector $\mathbf{x}$ tiene pendiente x_2/x_1, y la que contiene el vector $\mathbf{y}$ tiene pendiente y_2/y_1. Estas rectas son perpendiculares si y sólo si el producto de sus pendientes es igual a -1. Es decir, los vectores $\mathbf{x}$ y $\mathbf{y}$ serán perpendiculares si y sólo si

$$\left(\frac{x_2}{x_1}\right)\left(\frac{y_2}{y_1}\right) = -1$$

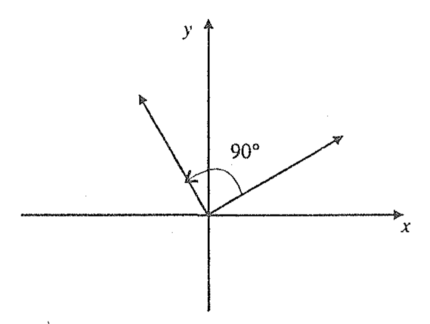


Figura 1. Dos vectores perpendiculares en el plano.

o sea

$$x_1y_1 + x_2y_2 = 0$$

El lado izquierdo de esta última expresión no es más que el producto punto de **x** y **y**. Podemos decir entonces que, en el plano $\mathbb{R}^2$, la perpendicularidad de vectores es equivalente al hecho de que su producto punto sea cero. Esta situación concreta motiva a establecer la siguiente definición, que generaliza la idea de perpendicularidad de vectores en el plano.

Definición. Se dice que los vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ son *ortogonales* si $\mathbf{x} \cdot \mathbf{y} = 0$.

Según esta definición, el vector $\mathbf{0} \in \mathbb{R}^n$ es ortogonal a cualquier vector $\mathbf{x} \in \mathbb{R}^n$, pues es claro que $\mathbf{0} \cdot \mathbf{y} = 0\ \forall \mathbf{y} \in \mathbb{R}^n$. Más aún, es el vector cero el *único* vector de $\mathbb{R}^n$ con esta propiedad (en efecto: si $\mathbf{x} \in \mathbb{R}^n$ es tal que $\mathbf{x} \cdot \mathbf{y} = 0\ \forall \mathbf{y} \in \mathbb{R}^n$, entonces, en particular se tiene $\mathbf{x} \cdot \mathbf{x} = 0$, por lo que, atendiendo a la primera propiedad del producto punto enunciada en el teorema 1.2.1, se concluye que $\mathbf{x}$ es el vector cero).

Ejemplo 1. Sea $\mathbf{u} = (1, -1)$. El vector $\mathbf{v} = (x, y)$ será ortogonal a $\mathbf{u}$ si y sólo si $\mathbf{u} \cdot \mathbf{v} = (1)(x) + (-1)(y) = x - y = 0$. Observe que el conjunto de todos los vectores v con esta propiedad constituyen geométricamente la recta $y = x$. Más en general, si $\mathbf{u} = (m, -1)$, donde m es un número real dado, los vectores $\mathbf{v} = (x, y)$ ortogonales a $\mathbf{u}$ son tales que $\mathbf{u} \cdot \mathbf{v} = mx - y = 0$. Estos vectores representan geométricamente la recta $y = mx$, la cual es una recta de pendiente m. Así pues, podemos decir que la recta $y = mx$ es el conjunto de todos los vectores en el plano ortogonales al vector $(m, -1)$. Generalmente, la recta $y = mx + b$, de pendiente m y ordenada al origen b, se puede describir como el conjunto de puntos (x, y) en el plano que se escriben como $(0, b) + (t, mt)$, con $t = x \in \mathbb{R}$, como se puede comprobar directamente. Esto significa que los vectores (x, y) de la recta $y = mx + b$ son vectores suma del vector constante $(0, b)$ con vectores del tipo (t, mt) que son ortogonales a $(m, -1)$ (es decir, estos vectores (t, mt) pertenecen a la recta $y = mx$).

Ejemplo 2. Dado el vector no nulo $\mathbf{u} = (a, b, c) \in \mathbb{R}^3$, el conjunto de vectores $\mathbf{v} \in \mathbb{R}^3$ ortogonales a $\mathbf{u}$ está formado por los vectores $\mathbf{v} = (x, y, z)$ de modo que $\mathbf{u} \cdot \mathbf{v} = ax + by + cz = 0$. La ecuación $ax + by + cz = 0$ representa geométricamente un plano que pasa por el origen. Este hecho será discutido con más amplitud en la sección 6 de este capítulo. ■

Un conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ en $\mathbb{R}^n$ se dice ser *ortogonal* si tomados dos a dos, éstos son ortogonales. Es decir, si $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, para todo $i \neq j,\ i, j = 1, 2, \ldots, n$.

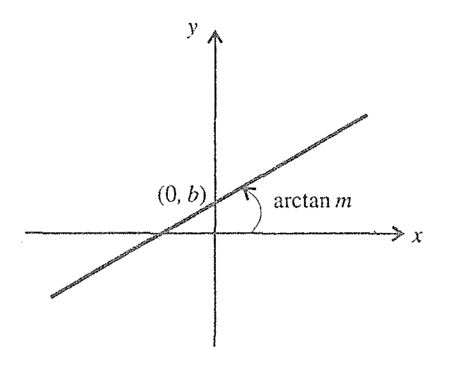


Figura 2. La recta $y = mx + b$.

Ejemplo 3. El conjunto $\{(-2, 4), (6, 3)\}$ es un conjunto ortogonal en $\mathbb{R}^2$, pues los dos únicos vectores que lo constituyen son ortogonales: $(-2, 4) \cdot (6, 3) = (-2)(6) + (4)(3) = 0$. Obsérvese que este conjunto ya "no acepta" otro vector no nulo (x, y); más precisamente, el conjunto $\{(-2, 4), (6, 3), (x, y)\}$ es ortogonal si sólo si (x, y) es el vector $\mathbf{0}$. En efecto, siendo (x, y) ortogonal a $(-2, 4)$ y a $(6, 3)$ se debe tener que $-2x + 4y = 0$, $4x + 3y = 0$, que es un sistema homogéneo de ecuaciones lineales que sólo tiene la solución $x = y = 0$. ■

Ejemplo 4. El conjunto $A = \{(-1, 1, 1), (2, 1, 1), (0, -1, 1)\}$ es un conjunto ortogonal de vectores en $\mathbb{R}^3$, pues

$$(-1, 1, 1) \cdot (2, 1, 1) = -2 + 1 + 1 = 0$$
$$(-1, 1, 1) \cdot (0, -1, 1) = 0 - 1 + 1 = 0$$
$$(2, 1, 1) \cdot (0, -1, 1) = 0 - 1 + 1 = 0$$

Se puede ver también que si (x, y, z) es un vector que unido a A produce un conjunto ortogonal, entonces éste es el vector cero. Esto es consecuencia de que el sistema homogéneo

$$-x + y + z = 0, \qquad 2x + y + z = 0, \qquad -y + z = 0$$

tiene solamente la solución trivial $x = y = z = 0$. ■

Los dos ejemplos anteriores nos hacen suponer que, en general, en el espacio $\mathbb{R}^n$ no se pueden tener conjuntos ortogonales de vectores no nulos con más de n vectores. Esto es cierto y es consecuencia del siguiente resultado.

Teorema 1.2.3 Sea $A = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ un conjunto de k vectores no nulos en $\mathbb{R}^n$. Si este conjunto es ortogonal, entonces es también linealmente independiente (y entonces necesariamente $k \leq n$).

Demostración. Debemos mostrar que de la combinación lineal

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = 0$$

se deduce que $c_1 = c_2 = \ldots c_k = 0$. Se tiene

$$0 = \mathbf{0} \cdot \mathbf{v}_i = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i = c_i(\mathbf{v}_i \cdot \mathbf{v}_i)$$

donde se usó la linealidad del producto punto y el hecho de que $\mathbf{v}_j \cdot \mathbf{v}_i = 0$ para toda j distinta de i. Puesto que $\mathbf{v}_i \cdot \mathbf{v}_i \neq \mathbf{0}$, concluimos que $c_i = 0$, lo cual es cierto para cualquier $i = 1, 2, \ldots, k$, como queríamos. Q.E.D.

Entonces, si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un conjunto ortogonal de vectores no nulos en $\mathbb{R}^n$, éste es linealmente independiente (y es, por tanto, una base del espacio). Si $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$ es otro vector tal que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$ sigue siendo un conjunto ortogonal, entonces, en particular $\mathbf{v}_i \cdot \mathbf{v} = 0$ para toda $i = 1, 2, \ldots, n$. Es decir que $0 = \mathbf{v}_i \cdot \mathbf{v} = \mathbf{v}_i \cdot (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n) = x_i(\mathbf{v}_i \cdot \mathbf{v}_i)$, de donde $x_i = 0$ (pues $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$, ¿por qué?). Así pues, el vector $\mathbf{v}$ es el vector cero.

Pasaremos ahora a discutir el importante concepto de proyección de un vector sobre otro. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores en $\mathbb{R}^n$ (para la discusión que presentamos a continuación, requerimos que $\mathbf{x}$ no sea ortogonal al vector $\mathbf{y}$). Tomemos la proyección ortogonal del vector $\mathbf{y}$ sobre el vector $\mathbf{x}$ como se muestra en la figura siguiente. Denotemos por $\mathbf{u}$ a este vector proyección (usaremos también la notación $\mathrm{PR}_{\mathbf{y}\to\mathbf{x}}$).

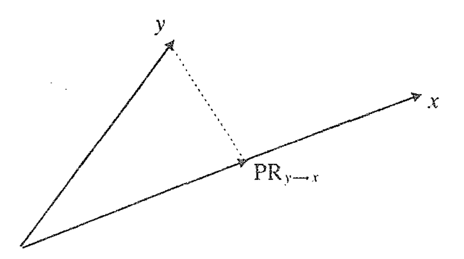


Figura 3. La proyección ortogonal del vector y sobre el vector x.

Es claro que el vector $\mathbf{u}$ es un múltiplo escalar del vector $\mathbf{x}$. Es decir, existe $\lambda \in \mathbb{R}$ tal que $\mathbf{u} = \lambda\mathbf{x}$. Observe además que el vector $\mathbf{v} = \mathbf{y} - \mathbf{u}$ es un vector ortogonal a $\mathbf{x}$. Entonces $(\mathbf{y} - \mathbf{u}) \cdot \mathbf{x} = \mathbf{0}$, o bien $(\mathbf{y} - \lambda\mathbf{x}) \cdot \mathbf{x} = \mathbf{0}$, de donde, usando las propiedades de linealidad del producto punto, obtenemos que

$$\lambda = \frac{\mathbf{y} \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}}$$

y así, la *proyección ortogonal de* $\mathbf{y}$ *sobre* $\mathbf{x}$ es el *vector*

$$\mathrm{PR}_{\mathbf{y}\to\mathbf{x}} = \frac{\mathbf{y} \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}}\mathbf{x}$$

Ejemplo 5. Si los vectores $\mathbf{x}$ y $\mathbf{y}$ son ortogonales, geométricamente es claro que $\mathrm{PR}_{\mathbf{y}\to\mathbf{x}}$ y $\mathrm{PR}_{\mathbf{x}\to\mathbf{y}}$ son iguales a (al vector) cero, lo cual se puede ver también de la fórmula para la proyección ortogonal, pues $\mathbf{x} \cdot \mathbf{y} = \mathbf{0}$. También, si $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ es un vector cualquiera de $\mathbb{R}^n$ y $\mathbf{e}_i$ es el i-ésimo vector de la base canónica de $\mathbb{R}^n$, entonces, puesto que $\mathbf{x} \cdot \mathbf{e}_i = x_i$ y $\mathbf{e}_i \cdot \mathbf{e}_i = 1$, se tiene

$$\mathrm{PR}_{\mathbf{x}\to\mathbf{e}_i} = \frac{\mathbf{x} \cdot \mathbf{e}_i}{\mathbf{e}_i \cdot \mathbf{e}_i}\mathbf{e}_i = \mathbf{x}_i\mathbf{e}_i$$

como era de esperarse. ■

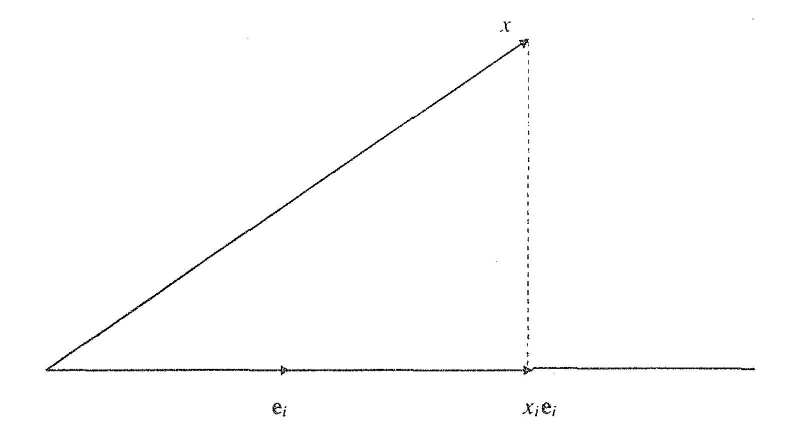


Figura 4. La proyección de $\mathbf{x}$ sobre $\mathbf{e}_i$ es $x_i\mathbf{e}_i$.

Ejemplo 6. Consideremos los vectores $x = (6, 2)$, $y = (1, 5)$ en $\mathbb{R}^2$. Se tiene

$$\mathrm{PR}_{\mathbf{y}\to\mathbf{x}} = \frac{\mathbf{y}\cdot\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}\mathbf{x} = \frac{(6)(1)+(2)(5)}{(6)^2+(2)^2}\mathbf{x} = \frac{2}{5}\mathbf{x}$$

$$\mathrm{PR}_{\mathbf{x}\to\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{x}}{\mathbf{y}\cdot\mathbf{y}}\mathbf{y} = \frac{(6)(1)+(2)(5)}{(1)^2+(5)^2}\mathbf{y} = \frac{8}{13}\mathbf{y}$$

lo cual se ve geométricamente como

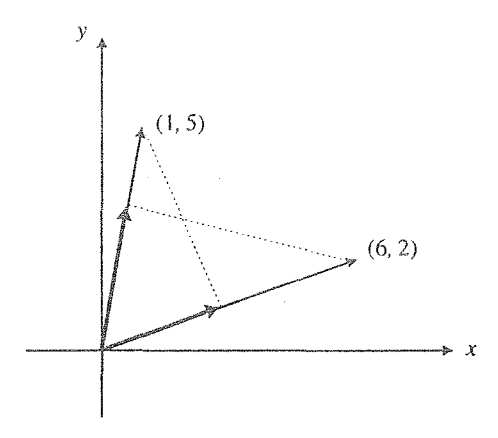


Figura 5. Los vectores del ejemplo 6.

Ejercicios (Capítulo 1, Sección 2)

1. Demuestre que $\mathbf{x}\cdot\mathbf{0} = 0$ $\forall\mathbf{x}\in\mathbb{R}^n$.
2. En este ejercicio se da un argumento distinto al del texto que prueba la validez de la desigualdad de Cauchy-Schwarz: $(\mathbf{x}\cdot\mathbf{y})^2 \leq (\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y})$. Sean $\mathbf{x}, \mathbf{y}$ dos vectores de $\mathbb{R}^n$. Si $\mathbf{y} = \mathbf{0}$, verifique

que se cumple automáticamente la desigualdad de Cauchy-Schwarz (de hecho, se cumple la igualdad). Si $\mathbf{y} \neq \mathbf{0}$, considere el vector $\mathbf{u} = \mathbf{x} - \alpha\mathbf{y}$, donde $\alpha = \frac{\mathbf{x}\cdot\mathbf{y}}{\mathbf{y}\cdot\mathbf{y}}$. Haga el producto punto de $\mathbf{u}$ consigo mismo y obtenga de ahí la desigualdad de Cauchy-Schwarz (usando que $\mathbf{u} \cdot \mathbf{u} \geq 0$). Use además el hecho de que $\mathbf{u} \cdot \mathbf{u} = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$, para concluir que la igualdad en la desigualdad de Cauchy-Schwarz se tiene si y sólo si los vectores $\mathbf{x}$, $\mathbf{y}$ son linealmente dependientes.

3. ¿Qué es el producto punto de dos vectores en $\mathbb{R}^1$? ¿Cómo se ve la desigualdad de Cauchy-Schwarz para vectores en $\mathbb{R}^1$? ¿En qué casos se tiene la igualdad en esta desigualdad para vectores en $\mathbb{R}^1$? (recuerde: el valor absoluto del producto de dos números reales es *igual* al producto de los valores absolutos de los números). Explique.

4. Verifique la desigualdad de Cauchy-Schwarz con los vectores

a. $\mathbf{x} = (1, 1)$, $\mathbf{y} = (3, 2)$

b. $\mathbf{x} = (2, 1, 1)$, $\mathbf{y} = (1, 0, 0)$

c. $\mathbf{x} = (3, 0, 1)$, $\mathbf{y} = (0, 0, 3)$

d. $\mathbf{x} = (1, 1, 1)$, $\mathbf{y} = (2, 2, 2)$

e. $\mathbf{x} = (1, 2, 0, 2, 1)$, $\mathbf{y} = (3, 1, 1, 0, 2)$

5. Sea $\mathbf{v}$ un vector de $\mathbb{R}^n$. Demuestre que el conjunto S de vectores $\mathbf{x} \in \mathbb{R}^n$ ortogonales a $\mathbf{v}$, es decir $S = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} \cdot \mathbf{v} = 0\}$ es un subespacio de $\mathbb{R}^n$.

6. Describa los vectores $(x, y) \in \mathbb{R}^2$ que son ortogonales al vector $(3, -1)$. Verifique que éstos son los puntos de una recta que pasa por el origen.

7. Describa los vectores $(x, y, z) \in \mathbb{R}^3$ que son ortogonales al vector $(-2, 1, 4)$. Verifique que éste es un subespacio de $\mathbb{R}^3$ del tipo

$$S = \{(x, y, z)|ax + by + cz = 0\}$$

Más en general, demuestre que todo subespacio S de $\mathbb{R}^3$ como el anterior, es descrito como $S = \{\mathbf{u} \in \mathbb{R}^3 | \mathbf{u} \cdot \mathbf{v} = 0\}$ para algún $\mathbf{v} \in \mathbb{R}^3$.

8. Escriba de manera vectorial cada una de las rectas dadas a continuación. Es decir, como un conjunto de vectores $(x, y) \in \mathbb{R}^2$ tales que $(x, y) = (0, b) + t(1, m)$, $t \in \mathbb{R}$, haciendo una gráfica en cada caso (ver ejemplo 1)

a. $y = 2x$

b. $y = x + 1$

c. $y = -2x + 3$

d. $y = -x - 1$

9. Considere la recta

$$\ell = \{(x, y)|(x, y) = (0, 3) + t(1, 2), t \in \mathbb{R}\}$$

Verifique que $(2, 7) \in \ell$. ¿Significa esto que el vector $(2, 7)$ se encuentra *sobre* la recta ℓ?. Explique.

10. a. Halle un vector $(x, y) \in \mathbb{R}^2$ que sea ortogonal a $(1, 2)$.

b. Halle un vector $(x, y) \in \mathbb{R}^2$ que sea ortogonal a $(1, 2)$ y a $(-3, -6)$.

c. Halle un vector $(x, y) \in \mathbb{R}^2$ que sea ortogonal a $(1, 2)$, $(-3, -6)$ y a $(2, 4)$.

d. Halle un vector $(x, y) \in \mathbb{R}^2$ que sea ortogonal a $(1, 2)$ y $(3, 5)$.

11. Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ dos vectores en $\mathbb{R}^2$ linealmente independientes. Demuestre que el único vector de $\mathbb{R}^2$ ortogonal a $\mathbf{v}_1$ y a $\mathbf{v}_2$ es el vector $\mathbf{0}$. ¿Ocurre lo mismo si los vectores no son linealmente independientes?

12. **a.** Halle un vector $(x, y, z) \in \mathbb{R}^3$ que sea ortogonal a $(3, 1, 1)$.
b. Halle un vector $(x, y, z) \in \mathbb{R}^3$ que sea ortogonal a $(3, 1, 1)$ y a $(6, 2, 2)$.
c. Halle un vector $(x, y, z) \in \mathbb{R}^3$ que sea ortogonal a $(3, 1, 1)$ y a $(2, 1, 5)$.
d. Halle un vector $(x, y, z) \in \mathbb{R}^3$ que sea ortogonal a $(3, 1, 1)$, $(2, 1, 5)$ y a $(1, 0, 0)$.

13. ¿Es cierta la afirmación del ejercicio 11 para vectores en el espacio $\mathbb{R}^3$?

14. Sean $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ tres vectores en $\mathbb{R}^3$ linealmente independientes. Demuestre que el único vector de $\mathbb{R}^3$ ortogonal a $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ es el vector $\mathbf{0}$. ¿Ocurre lo mismo si los vectores no son linealmente independientes?

15. Sea $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base de $\mathbb{R}^n$. Demuestre que el único vector $\mathbf{v} \in \mathbb{R}^n$ ortogonal a todos y cada uno de los vectores de β es el vector $\mathbf{0}$.

16. Del teorema 1.2.3 se dedujo que todo conjunto ortogonal de n vectores no nulos en $\mathbb{R}^n$ es una base de este espacio. ¿Es cierta la afirmación recíproca?

17. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ un conjunto ortogonal de k vectores en $\mathbb{R}^n$. Demuestre que si $k > n$, entonces alguno de los vectores de este conjunto es el vector $\mathbf{0}$.

18. ¿Cierto o falso? Todo conjunto de vectores que contenga al vector $\mathbf{0}$ es ortogonal.

19. ¿Cierto o falso? Un conjunto con un solo vector no puede ser ortogonal.

20. Considere el cuadrilátero cuyos vértices son $A = (1, -2, 2)$, $B = (1, 4, 0)$, $C = (-4, 1, 1)$, $D = (-5, -5, 3)$. Demuestre que las diagonales AC y BD son ortogonales.

21. Verifique que los siguientes conjuntos de vectores son conjuntos ortogonales
a. $\{(3, 1), (2, -6)\}$
b. $\{(a, 0), (0, b)\}$ en donde a y b son números dados.
c. $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$
d. $\{(4, -1, 0), (2, 3, -5), (-8, 7, 1)\}$
e. $\{(2, 0, 1, 3), (-2, 4, 1, 1), (-3, -2, 0, 2)\}$

22. Demostrar que los puntos $A = (1, 1)$, $B = (2, 3)$ y $C = (5, -1)$ son los vértices de un triángulo rectángulo.

23. En cada uno de los incisos siguientes, encuentre el vector $\mathbf{u} = \mathrm{PR}_{\mathbf{y} \to \mathbf{x}}$, proyección del vector y sobre el vector $\mathbf{x}$. Verifique en cada caso que el vector obtenido es ortogonal a $\mathbf{y} - \mathbf{u}$.
a. $x = (2, 5)$, $y = (3, 4)$
b. $x = (1, 0)$, $y = (4, 5)$
c. $x = (4, 2)$, $y = (2, 1)$
d. $x = (2, 1, 0)$, $y = (1, 0, 1)$
e. $x = (1, 1, 1, 2)$, $y = (0, 2, 0, 3)$

24. Sea $\mathbf{v}$ el vector de $\mathbb{R}^3$ cuyo punto inicial está en $(1, 3, 7)$ y cuyo punto final está en $(4, 5, 7)$. Hallar la proyección del vector $(1, 2, 1)$ sobre $\mathbf{v}$.

25. Use el concepto de proyección de un vector sobre otro para calcular el área del triángulo cuyos vértices son:

a. $A = (0, 0)$, $B = (5, 3)$, $C = (7, 8)$

b. $A = (0, 0)$, $B = (9, 1)$, $C = (5, 4)$

c. $A = (-2, -3)$, $B = (3, 2)$, $C = (-1, 5)$

d. $A = (1, 3, 2)$, $B = (2, 5, 3)$, $C = (-2, 0, 0)$

26. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores en $\mathbb{R}^n$ y $-\mathbf{x}$, $-\mathbf{y}$ sus inversos aditivos. Demuestre que

a. $\mathrm{PR}_{\mathbf{y}\to-\mathbf{x}} = \mathrm{PR}_{\mathbf{y}\to\mathbf{x}}$

b. $\mathrm{PR}_{-\mathbf{y}\to\mathbf{x}} = -\mathrm{PR}_{\mathbf{y}\to\mathbf{x}}$

Verifique este resultado con los vectores $\mathbf{x} = (2, 5)$, $\mathbf{y} = (-1, 3)$.

27. Sean $\mathbf{v}, \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$, $k + 1$ vectores en $\mathbb{R}^n$. Si $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k$. Demuestre que

$$\mathrm{PR}_{\mathbf{u}\to\mathbf{v}} = \mathrm{PR}_{\mathbf{u}_1\to\mathbf{v}} + \mathrm{PR}_{\mathbf{u}_2\to\mathbf{v}} + \cdots + \mathrm{PR}_{\mathbf{u}_k\to\mathbf{v}}$$

Verifique este resultado con los vectores $\mathbf{u}_1 = (1, 1)$, $\mathbf{u}_2 = (3, -2)$, $\mathbf{v} = (2, 3)$.

28. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores en $\mathbb{R}^n$ y c un número real. Demuestre que

$$\mathrm{PR}_{c\mathbf{u}\to\mathbf{v}} = c\,\mathrm{PR}_{\mathbf{u}\to\mathbf{v}}$$

Verifique este resultado con $\mathbf{u} = (2, 1, -1)$, $\mathbf{v} = (3, 5, 1)$, $c = 2$.

1.3 Norma y distancia

Con la ayuda del producto punto estudiado en la sección anterior, y con el cual ya tenemos en $\mathbb{R}^n$ el concepto geométrico de ortogonalidad de vectores, es posible introducir una noción de "tamaño de un vector" y de "distancia entre dos vectores" (o distancia entre dos puntos en $\mathbb{R}^n$).

Definimos la *norma* (de modo más preciso, la *norma euclidiana*) de un vector $\mathbf{x} \in \mathbb{R}^n$, denotada por $\|\mathbf{x}\|$, como

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

(según la primera propiedad del producto punto de vectores, esta definición hace perfecto sentido, pues lo que está dentro de la raíz cuadrada es siempre un número no negativo). En concreto, si $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, se tiene

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Diremos que el vector $\mathbf{x}$ es *unitario* si $\|\mathbf{x}\| = 1$. Obsérvese que los vectores de la base canónica de $\mathbb{R}^n$ son unitarios, pues

$$\|\mathbf{e}_i\| = \sqrt{(0)^2 + \cdots + (1)^2 + \cdots + (0)^2} = 1, \qquad i = 1, 2, \ldots, n$$

Viendo los casos de vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$, debe ser claro que esta noción de norma de un vector, que ha sido definida en general en el espacio $\mathbb{R}^n$, nos da una medida del "tamaño" del vector. En efecto,

si $\mathbf{v} = (x, y) \in \mathbb{R}^2$, se tiene $\|\mathbf{v}\| = \sqrt{x^2 + y^2}$, que no es más que la distancia del punto (x, y) al origen (i.e. es justamente el tamaño del vector $\mathbf{v}$). Una situación análoga se puede ver en el espacio $\mathbb{R}^3$.

En la sección anterior obtuvimos la desigualdad de Cauchy-Schwarz

$$(\mathbf{x} \cdot \mathbf{y})^2 \leq (\mathbf{x} \cdot \mathbf{x})(\mathbf{y} \cdot \mathbf{y})$$

Si tomamos raíz cuadrada en ambos miembros de esta desigualdad, nos queda

$$|\mathbf{x} \cdot \mathbf{y}| \leq \sqrt{\mathbf{x} \cdot \mathbf{x}}\sqrt{\mathbf{y} \cdot \mathbf{y}}$$

que podemos escribir, usando el concepto de norma como

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

Esta "versión" de la desigualdad de Cauchy-Schwarz es la que usaremos cuando se requiera.

En el teorema siguiente se recogen las principales propiedades de la norma de un vector en $\mathbb{R}^n$.

Teorema 1.3.1 Sean $\mathbf{x}, \mathbf{y}$ vectores de $\mathbb{R}^n$, y c un número real. Se tiene

a. $\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

b. $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$

c. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (desigualdad triangular)

Demostración. El hecho de que $\|\mathbf{x}\| \geq 0$ es inmediato de la definición. También, el hecho que $\|\mathbf{x}\| = 0$, equivale a que $\mathbf{x} \cdot \mathbf{x} = 0$, lo cual a su vez equivale (según la primera propiedad del producto punto) a que $\mathbf{x} = \mathbf{0}$. La propiedad b) se obtiene madiante operaciones sencillas:

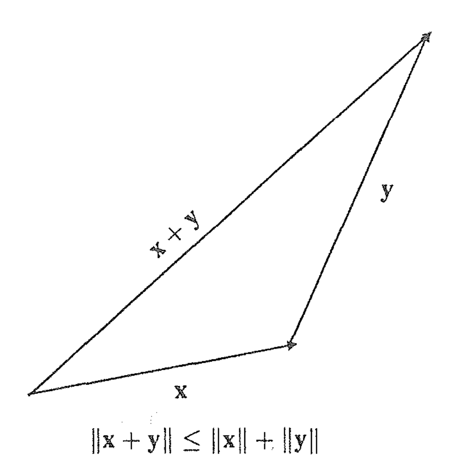
$$\|c\mathbf{x}\| = \sqrt{c\mathbf{x} \cdot c\mathbf{x}} = \sqrt{c^2(\mathbf{x} \cdot \mathbf{x})} = \sqrt{c^2}\sqrt{\mathbf{x} \cdot \mathbf{x}} = |c|\|\mathbf{x}\|$$

Para ver la validez de la propiedad c), escribimos

$$\begin{aligned}
\|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) && \begin{pmatrix}\text{definición} \\ \text{de norma}\end{pmatrix} \\
&= \mathbf{x} \cdot \mathbf{x} + 2(\mathbf{x} \cdot \mathbf{y}) + \mathbf{y} \cdot \mathbf{y} && \begin{pmatrix}\text{bilinealidad del} \\ \text{producto punto}\end{pmatrix} \\
&\leq \|\mathbf{x}\|^2 + 2|\mathbf{x} \cdot \mathbf{y}| + \|\mathbf{y}\|^2 && \begin{pmatrix}\text{definición de} \\ \text{norma, } \alpha \leq |\alpha|\end{pmatrix} \\
&\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 && \begin{pmatrix}\text{desigualdad de} \\ \text{Cauchy-Schwarz}\end{pmatrix} \\
&= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2
\end{aligned}$$

de donde se obtiene la desigualdad procurada (tomando raíz cuadrada en ambos miembros de la desigualdad obtenida). Q.E.D.

La propiedad c) del teorema anterior describe un hecho geométrico bien conocido: en un triángulo cualquiera, la longitud de uno de sus lados no puede exceder a la suma de las longitudes de los otros dos lados. Es por eso que el resultado es conocido como *desigualdad triangular*. La siguiente figura aclara esta situación.



$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Figura 1. La desigualdad triangular.

Ejemplo 1. Verifiquemos la desigualdad triangular con los vectores $\mathbf{x} = (1, -2, 3, 2)$, $\mathbf{y} = (5, -3, 0, 1)$. Se tiene que $\mathbf{x}+\mathbf{y} = (6, -5, 3, 3)$, de modo que

$$\|\mathbf{x}+\mathbf{y}\| = \|(6, -5, 3, 3)\| = \sqrt{79}$$

Por otra parte

$$\|\mathbf{x}\| = \|(1, -2, 3, 2)\| = \sqrt{18} \qquad \|\mathbf{y}\| = \|(5, -3, 0, 1)\| = \sqrt{35}$$

Se tiene efectivamente que

$$\sqrt{79} = \|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| = \sqrt{18} + \sqrt{35}$$ ■

Con la ayuda del concepto de norma (y lo estudiado en la sección anterior), podemos introducir fácilmente el concepto de ángulo entre dos vectores en $\mathbb{R}^n$. En el caso de dos vectores en $\mathbb{R}^2$, es fácil obtener una expresión para el ángulo que forman. En efecto, sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ dos vectores no nulos. De la figura 3 es inmediato que el ángulo θ que forman $\mathbf{x}$ y $\mathbf{y}$ es tal que

$$\cos\theta = \frac{\|\operatorname{PR}_{\mathbf{y}\to\mathbf{x}}\|}{\|\mathbf{y}\|} = \frac{\frac{|\mathbf{x}\cdot\mathbf{y}|}{\|\mathbf{x}\|}}{\|\mathbf{y}\|} = \frac{|\mathbf{x}\cdot\mathbf{y}|}{\|\mathbf{x}\|\|\mathbf{y}\|}$$

La fórmula anterior tiene sentido si nuestros vectores $\mathbf{x}, \mathbf{y}$ son vectores cualesquiera no nulos del espacio $\mathbb{R}^n$. De hecho, se define el *ángulo entre los vectores* (no nulos) $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ como el ángulo θ, $0 \leq \theta \leq \pi$, dado por

$$\theta = \arccos\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$$

Obsérvese que si los vectores $\mathbf{x}, \mathbf{y}$ son ortogonales, entonces el ángulo que forman entre ellos es $\theta = \arccos 0 = \pi/2$, como era de esperarse. Hemos quitado el valor absoluto en el escalar $\mathbf{x}\cdot\mathbf{y}$ de la expresión obtenida para θ de la figura 2. Esto lo hacemos para dejar la posibilidad de ángulos *obtusos* entre los vectores $\mathbf{x}$ y $\mathbf{y}$. De hecho, es claro que si $\mathbf{x}\cdot\mathbf{y}$ es negativo, el ángulo θ será obtuso; si $\mathbf{x}\cdot\mathbf{y}$ es positivo, el ángulo θ será agudo (y, como ya se dijo, si $\mathbf{x}\cdot\mathbf{y} = 0$, el ángulo θ es recto).

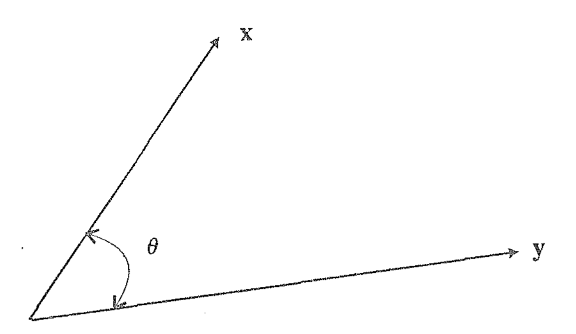


Figura 2. El ángulo θ entre los vectores $\mathbf{x}$, $\mathbf{y}$.

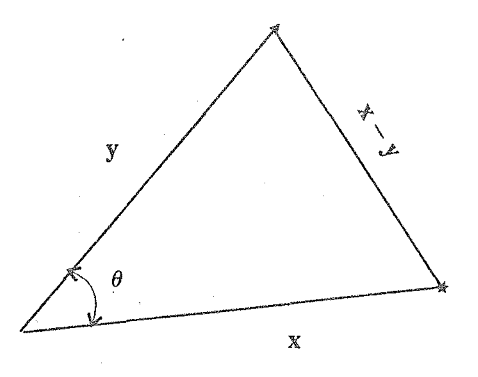
Nótese también que en términos del ángulo θ, se puede escribir el producto punto de los vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ como

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$$

Más aún, si calculamos directamente el cuadrado de la norma del vector diferencia $\mathbf{x} - \mathbf{y}$, obtenemos

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \\ &= \mathbf{x} \cdot \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{y}) + \mathbf{y} \cdot \mathbf{y} \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\theta \end{aligned}$$

la cual no es más que la versión (¡generalizada!) para vectores en $\mathbb{R}^n$ de la conocida "ley de los cosenos" que se estudia en los cursos de trigonometría elemental.



$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$$

Figura 3. La ley de los cosenos.

Ejemplo 2. El ángulo entre los vectores $\mathbf{x} = (1, -2, 3, 2)$, $\mathbf{y} = (3, 4, 0, 8)$ es

$$\theta = \arccos \frac{x \cdot y}{\|x\|\|y\|} = \arccos \frac{11}{\sqrt{18}\sqrt{89}} = \arccos \frac{11}{3\sqrt{178}}$$

que aproximadamente es de 74 grados. ■

Debemos mencionar que el concepto de norma es más general que el presentado anteriormente: una *norma* en el espacio $\mathbb{R}^n$ es una función $\| \ \|: \mathbb{R}^n \to \mathbb{R}$ que a cada vector $\mathbf{x} \in \mathbb{R}^n$ le asocia el número real $\|\mathbf{x}\|$(la norma de $\mathbf{x}$) y que cumple con las tres propiedades establecidas en el teorema 1.3.1. La norma que presentamos aquí es la llamada "norma euclidiana" y con ella trabajaremos en este libro. Existen, sin embargo, otras normas importantes en $\mathbb{R}^n$ (digamos que otras maneras de medir el tamaño de los vectores en $\mathbb{R}^n$), por ejemplo

(*) la norma del máximo $\| \ \|_{\text{máx}}: \mathbb{R}^n \to \mathbb{R}$ dada por

$$\|\mathbf{x}\|_{\text{máx}} = \text{máx}(|x_i|, i = 1, 2, \ldots, n)$$

(*) la norma de la suma $\| \ \|_s: \mathbb{R}^n \to \mathbb{R}$ dada por

$$\|\mathbf{x}\|_s = |x_1| + |x_2| + \cdots + |x_n|$$

donde $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Dejamos al lector que verifique que éstas son efectivamente normas en $\mathbb{R}^n$.

Estudiemos ahora el concepto de distancia entre dos vectores en $\mathbb{R}^n$. Recordemos que el vector diferencia $\mathbf{x} - \mathbf{y}$ de $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ es un vector que "conecta los puntos finales de las flechas que representan a $\mathbf{x}$ y $\mathbf{y}$". La norma de este vector es entonces una medida de la distancia que separa a los *puntos* $\mathbf{x}$ y $\mathbf{y}$ en el espacio $\mathbb{R}^n$. Esta es, de hecho, la definición que daremos de distancia entre dos vectores en $\mathbb{R}^n$. Dados $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, definimos la *distancia* entre $\mathbf{x}$ y $\mathbf{y}$, denotada por $d(\mathbf{x}, \mathbf{y})$, como

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

Esquemáticamente se tiene

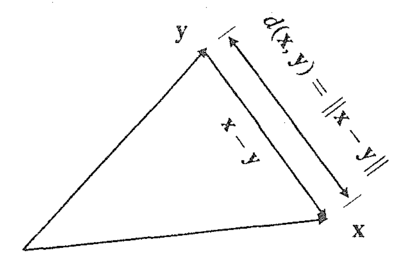


Figura 4. Distancia entre x y y.

Haciendo explícitas las coordenadas de los vectores x y y, poniendo $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ se tiene

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_3)^2 + \cdots + (x_n - y_n)^2}$$

Nuevamente, los casos $n = 2$ y $n = 3$, nos dan las conocidas fórmulas de la "distancia entre dos puntos" en el plano y en el espacio, respectivamente, estudiadas en los cursos de geometría analítica: la distancia entre el punto $p_1 = (x_1, y_1)$ y $p_2 = (x_2, y_2)$ es

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Análogamente, si $\mathbf{p}_1 = (x_1, y_1, z_1)$, $\mathbf{p}_2 = (x_2, y_2, z_2)$ son dos puntos en $\mathbb{R}^3$, la distancia entre ellos es

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

El siguiente teorema recoge las propiedades más importantes de la distancia entre dos puntos en $\mathbb{R}^n$.

Teorema 1.3.2 Sean x, y dos vectores cualesquiera en $\mathbb{R}^n$. Se tiene

a. $d(\mathbf{x}, \mathbf{y}) \geq 0$, $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$.

b. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$

c. $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$, z un vector cualquiera de $\mathbb{R}^n$.

Demostración. El hecho de que $d(\mathbf{x}, \mathbf{y})$ sea un número no negativo, es consecuencia de la definición misma y de la primera propiedad de la norma establecida en el teorema 1.3.1. Además, por esta misma propiedad se tiene que $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = 0$ si y sólo si $\mathbf{x} - \mathbf{y} = 0$, o sea si y sólo si $\mathbf{x} = \mathbf{y}$ como se quería. Para ver la segunda propiedad hacemos uso de la propiedad b) de la norma, quedándonos

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \|(-1)(\mathbf{y} - \mathbf{x})\| = |-1|\|\mathbf{y} - \mathbf{x}\| = \|\mathbf{y} - \mathbf{x}\| = d(\mathbf{y}, \mathbf{x})$$

Por último, la propiedad c) no es más que una versión un poco más general de la desigualdad triangular ya demostrada en el teorema 1.3.1, pues

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &= \|\mathbf{x} - \mathbf{y}\| \\ &= \|(\mathbf{x} - \mathbf{z}) + (\mathbf{z} - \mathbf{y})\| \\ &\leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\| \\ &= d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}) \end{aligned}$$

Q.E.D

Ejemplo 3. La distancia entre el vector $\mathbf{x} = (2, 3, 3, 4, 1)$ y el vector $\mathbf{y} = (1, 1, 2, -1, 3)$ es

$$d = \|(2, 3, 3, 4, 1) - (1, 1, 2, -1, 3)\| = \|(1, 2, 1, 5, -2)\| = \sqrt{35}$$

■

Para terminar esta sección, hacemos nuevamente hincapié en que el concepto de distancia entre dos vectores de $\mathbb{R}^n$ es más general (la misma situación que ocurre con la norma) y que el que aquí presentamos es un caso particular de *métrica* en este espacio (llamada métrica o distancia euclidiana):

una métrica en $\mathbb{R}^n$ es una función $d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, que asocia a cada par de vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ un número real $d(\mathbf{x}, \mathbf{y})$ (llamado "distancia entre $\mathbf{x}$ y $\mathbf{y}$") y que satisface las tres propiedades establecidas en el teorema anterior. Obsérvese que estas propiedades representan lo que cualquier concepto sensato de distancia entre dos puntos debería cumplir: la distancia entre dos puntos siempre es un número no negativo, y es cero en el caso (y sólo en el caso) en que se esté midiendo la distancia de un vector a él mismo; la distancia entre $\mathbf{x}$ y $\mathbf{y}$ es la misma que entre $\mathbf{y}$ y $\mathbf{x}$; por último, se pide que se cumpla la desigualdad triangular (para asegurar que se está midiendo adecuadamente distancias relativas entre tres puntos). Algunas otras métricas en $\mathbb{R}^n$ son:

(*) La métrica discreta $d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $d(\mathbf{x}, \mathbf{y}) = 1$ si $\mathbf{x} \neq \mathbf{y}$, $d(\mathbf{x}, \mathbf{y}) = 0$ si $\mathbf{x} = \mathbf{y}$. (Este es un ejemplo poco importante, pero interesante: nos permite entender que el concepto de métrica establecido anteriormente... ¡es realmente muy general!).

(*) La métrica $d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $d(\mathbf{x}, \mathbf{y}) = \max(|x_i - y_i|, i = 1, 2, \ldots n)$.

(*) La métrica $d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$.

en donde $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. Dejamos a cargo del lector la verificación de que los tres ejemplos anteriores son efectivamente métricas en $\mathbb{R}^n$.

Ejercicios (Capítulo 1, Sección 3)

1. Calcule la norma de los siguientes vectores

   a. $(4, 0)$
   b. $(-3, 1)$
   c. $(1, 2, 1)$
   d. $(2, -3, 4)$
   e. $(2, 0, 2)$
   f. $(1, 3, 2, 0, -1)$

2. ¿Es la función norma $\|\cdot\|: \mathbb{R}^n \to \mathbb{R}$, que a cada vector $\mathbf{x} \in \mathbb{R}^n$ le asocia su norma $\|\mathbf{x}\| \in \mathbb{R}$, una función inyectiva?, ¿sobreyectiva?.

3. a. Sean a y b dos números reales no nulos. Demuestre que los cuatro vectores $(\pm a, \pm b) \in \mathbb{R}^2$ tienen la misma norma. Interprete geométricamente este hecho.

   b. Sean a, b y c tres números reales no nulos. Demuestre que los 8 vectores $(\pm a, \pm b, \pm c) \in \mathbb{R}^3$ tienen la misma norma. Intereprete geométricamente este hecho.

   c. Sean $x_1, x_2, \ldots, x_n$ n números reales no nulos. Demuestre que los 2^n vectores $(\pm x_1, \pm x_2, \ldots, \pm x_n) \in \mathbb{R}^n$ tienen la misma norma.

4. ¿Qué es la norma de un vector $\mathbf{x}$ en $\mathbb{R}^1$? ¿Cómo se ven las propiedades de la norma (teorema 1.3.1) en este caso?

5. Use la desigualdad de Cauchy-Schwarz para probar que si $x_1, x_2, \ldots, x_n$ son números reales cualesquiera, entonces

$$\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n} x_i\right| \leq \sqrt{\sum_{i=1}^{n} x_i^2}$$

   y que la igualdad se da si y sólo si todos los x_i son iguales.

6. Verifique la desigualdad triangular con los vectores

a. $\mathbf{x} = (0, 5)$, $\mathbf{y} = (1, 7)$

b. $\mathbf{x} = (2, 1, -2)$, $\mathbf{y} = (3, 1, 2)$

c. $\mathbf{x} = (-2, 0, 2, 1)$, $\mathbf{y} = (0, 0, 3, 7)$

7. Demuestre que la igualdad en la desigualdad triangular se tiene si y sólo si uno de los vectores es un múltiplo *no negativo* del otro vector.

8. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores en $\mathbb{R}^n$. Demuestre que

$$| \|\mathbf{x}\| - \|\mathbf{y}\| | \leq \|\mathbf{x} - \mathbf{y}\|$$

(Sugerencia: $\|\mathbf{x}\| = \|(\mathbf{x}-\mathbf{y})+\mathbf{y}\| \leq \|\mathbf{x}-\mathbf{y}\|+\|\mathbf{y}\|$, de donde se obtiene que $\|\mathbf{x}\|-\|\mathbf{y}\| \leq \|\mathbf{x}-\mathbf{y}\|$. De la misma manera se obtiene que $\|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\|$).

9. ¿Existen vectores $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ tales que $\|\mathbf{x}\| = 3$, $\|\mathbf{y}\| = 1$, $\|\mathbf{x} + \mathbf{y}\| = 5$? ¿Existen vectores $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ tales que $\|\mathbf{x}\| = 3$, $\|\mathbf{y}\| = 1$, $\|\mathbf{x} - \mathbf{y}\| = 5$?

10. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores ortogonales en $\mathbb{R}^n$, tales que $\|\mathbf{x}\| = 3$, $\|\mathbf{y}\| = 7$. Calcule $\|\mathbf{x}+\mathbf{y}\|$, $\|\mathbf{x}-\mathbf{y}\|$.

11. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores en $\mathbb{R}^n$ tales que $\|\mathbf{x}\| = 4$, $\|\mathbf{y}\| = 5$, $\|\mathbf{x} + \mathbf{y}\| = 7$. Calcule $\|\mathbf{x} - \mathbf{y}\|$.

12. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores en $\mathbb{R}^n$ tales que $\|\mathbf{x}\| = 11$, $\|\mathbf{y}\| = 23$, $\|\mathbf{x} - \mathbf{y}\| = 30$. Calcule $\|\mathbf{x} + \mathbf{y}\|$.

13. Demuestre que si $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ es un conjunto ortogonal de vectores en $\mathbb{R}^n$, entonces

$$\|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 + \cdots + \|\mathbf{x}_k\|^2$$

A este resultado se le conoce como "Teorema de Pitágoras". ¿Qué tiene que ver con el resultado del mismo nombre que se estudia en la matemática elemental?

14. Sean $\mathbf{x}$, $\mathbf{y}$ dos vectores en $\mathbb{R}^n$. Demuestre que

a. $\mathbf{x} \cdot \mathbf{y} \in \mathbb{R}^+$ si y sólo si $\|\mathbf{x} + \mathbf{y}\| > \|\mathbf{x} - \mathbf{y}\|$.

b. $-(\mathbf{x} \cdot \mathbf{y}) \in \mathbb{R}^+$ si y sólo si $\|\mathbf{x} + \mathbf{y}\| < \|\mathbf{x} - \mathbf{y}\|$.

c. $\mathbf{x}$ e $\mathbf{y}$ son ortogonales si y sólo si $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|$.

Discuta el contenido geométrico de estos resultados.

15. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores en $\mathbb{R}^n$ tales que $\|\mathbf{u}\| = \|\mathbf{v}\|$. Demuestre que los vectores $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ son ortogonales. ¿Vale la afirmación recíproca?

16. Calcule el ángulo entre los vectores del ejercicio 6.

17. Calcule el ángulo entre un vector $\mathbf{v} \in \mathbb{R}^n$ no nulo y su inverso aditivo.

18. Los vectores $\mathbf{u}$ y $\mathbf{v}$ de $\mathbb{R}^n$ forman un ángulo de $\pi/3$. Suponiendo que $\|\mathbf{u}\| = 3$ y $\|\mathbf{v}\| = 4$, calcule: $\mathbf{u} \cdot \mathbf{v}$; $\|\mathbf{u} + \mathbf{v}\|$; $\|\mathbf{u} - \mathbf{v}\|$

19. Cada pareja de vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en $\mathbb{R}^n$ forma un ángulo de $\pi/3$. Suponga que $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$, $\|\mathbf{w}\| = 3$. Calcule $\|\mathbf{u} + \mathbf{v} + \mathbf{w}\|$.

20. Sea $\{\mathbf{u}, \mathbf{v}\}$ un conjunto ortogonal de vectores unitarios en $\mathbb{R}^n$. Demuestre que el ángulo entre el vector $\mathbf{u}$ y el vector $\mathbf{u} + \mathbf{v}$ es de $\pi/4$. Discuta el contenido geométrico de este resultado cuando $n = 2$. (Sugerencia: use el teorema de Pitágoras para calcular $\|\mathbf{u} + \mathbf{v}\|$).

21. ¿Vale el resultado del ejercicio anterior si los vectores $\mathbf{u}$ y $\mathbf{v}$ no son unitarios?

22. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores en $\mathbb{R}^n$. Demuestre que

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2)$$

A este resultado se le conoce como "Ley del Paralelogramo". Justifique este nombre en base a su contenido geométrico en el caso $n = 2$. Resuelva de nuevo los ejercicios 11 y 12 a la luz de este resultado.

23. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores no nulos en $\mathbb{R}^n$ tales que $\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{u}-\mathbf{v}\|$. Demuestre que el ángulo entre $\mathbf{u}$ y $\mathbf{v}$ es de $\pi/3$. ¿Cuál es el ángulo entre $\mathbf{u}$ y $\mathbf{u}-\mathbf{v}$?, ¿y entre $\mathbf{v}$ y $\mathbf{u}-\mathbf{v}$? Discuta el contenido geométrico de este ejercicio en el caso $n = 2$.

24. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores no nulos en $\mathbb{R}^n$ tales que $\|\mathbf{u}\| = \|\mathbf{u}-\mathbf{v}\|$. Demuestre que el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$ es el mismo que el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{u}-\mathbf{v}$. Discuta el contenido geométrico de este ejercicio en el caso $n = 2$.

25. Considere las rectas

$$\ell_1 = \{(x, y) | (x, y) = (0, b_1) + t(1, m_1), \quad t \in \mathbb{R}\}$$
$$\ell_2 = \{(x, y) | (x, y) = (0, b_2) + t(1, m_2), t \in \mathbb{R}\}$$

(ver ejemplo 1 de la sección anterior). Con los vectores $\mathbf{v}_1 = (1, m_1)$ y $\mathbf{v}_2 = (1, m_2)$ que son paralelos a ℓ_1 y ℓ_2 respectivamente, demuestre, partiendo de la fórmula para el ángulo entre dos vectores, que el ángulo θ $(0 \le \theta \le \pi)$ entre ℓ_1 y ℓ_2 es

$$\theta = \arctan \frac{m_2 - m_1}{1 + m_2 m_1}$$

26. Sea $\mathbf{u} = (6, -8, -15/2)$. Determine el vector $\mathbf{v} \in \mathbb{R}^3$ sabiendo que es linealmente dependiente con $\mathbf{u}$, que $\|\mathbf{v}\| = 50$, y que el ángulo que forma $\mathbf{v}$ con la parte positiva del eje z es agudo.

27. Calcule la distancia entre cada par de los vectores siguientes

a. $\mathbf{x} = (7, 1)$, $\mathbf{y} = (3, 5)$.

b. $\mathbf{x} = (3, 4, 1)$, $\mathbf{y} = (2, 1, 1)$.

c. $\mathbf{x} = (2, 1, 1, 1)$, $\mathbf{y} = (1, 0, 0, 4)$.

28. Demuestre que el triángulo cuyos vértices son $A = (1, 1)$, $B = (4, 3)$ y $C = (1/2, 5)$ es isósceles. Determine sus ángulos internos.

29. Repita el ejercicio anterior con los puntos $A = (1, 2, 1)$, $B = (3, -1, 7)$, $C = (7, 4, -2)$.

30. Demostrar que los puntos $A = (2, 2)$, $B = (-1, 6)$, $C = (-5, 3)$ y $D = (-2, -1)$ son los vértices de un cuadrado.

31. Sean $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ dos vectores en $\mathbb{R}^n$. Demuestre que el punto $\mathbf{p} = \frac{1}{2}(\mathbf{x}+\mathbf{y})$ es un punto equidistante de $\mathbf{x}$ y $\mathbf{y}$ (es decir, $d(\mathbf{x}, \mathbf{p}) = d(\mathbf{y}, \mathbf{p})$), el cual se encuentra "sobre el segmento que une a $\mathbf{x}$ con $\mathbf{y}$" (para ver esto, nótese que los vectores $\mathbf{u} = \mathbf{x} - \mathbf{p}$, $\mathbf{v} = \mathbf{y} - \mathbf{p}$, son linealmente dependientes). Se dice que $\mathbf{p}$ es el punto medio del segmento $\overline{\mathbf{xy}}$. Determine el punto medio del segmento $\overline{\mathbf{xy}}$ en cada uno de los siguientes casos.

a. $\mathbf{x} = (2, 5)$, $\mathbf{y} = (8, 15)$

b. $\mathbf{x} = (3, 6, 9)$, $\mathbf{y} = (3, 3, -7)$

c. $\mathbf{x} = (1, 1, 1, 1)$, $\mathbf{y} = (3, 3, 3, 3)$

32. Los vértices de un triángulo son $A = (2, 4)$, $B = (6, 6)$ y $C = (3, 7)$. Determinar las coordenadas de los puntos medios de sus lados.

33. Los puntos medios de los lados de un triángulo son $P = (2, -1)$, $Q = (-1, 4)$ y $R = (-2, 2)$. Determinar los vértices del triángulo.

34. Determinar la longitud de la mediana del triángulo del ejercicio 32 trazada por el vértice B. (Recuerde que la mediana por el vértice V es la recta que va de V al punto medio del lado opuesto de V).

35. Use el concepto de proyección de un vector sobre otro para demostrar que la distancia del punto $\mathbf{p} = (x_0, y_0)$ a la recta $Ax + By + C = 0$ viene dada por la fórmula

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}$$

36. En cada uno de los incisos siguientes, calcule la distancia entre las dos líneas paralelas dadas:

a. $3x - 4y + 5 = 0$, $3x - 4y - 5 = 0$

b. $x + y = 0$, $x + y = -3$

c. $2x - y - 5 = 0$, $4x - 2y - 10 = 0$

d. $x + 4y - 2 = 0$, $-2x - 8y + 7 = 0$

37. Considere la recta $Ax + By - bB = 0$, en donde b es la ordenada al origen. Demuestre que las rectas paralelas que se encuentran a d unidades de la recta dada son $Ax + By - bB \pm \sqrt{A^2 + B^2}d = 0$.

38. Considere la recta $Ax + By + C = 0$, la cual dista r unidades del origen. Demuestre que la recta paralela a la recta dada, que dista también r unidades del origen (y que resulta ser simétrica respecto del origen de la recta dada), es $Ax + By - C = 0$.

39. **(Cubos de $\mathbb{R}^n$).** Sea $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ la base canónica de $\mathbb{R}^n$. Al conjunto $C \subset \mathbb{R}^n$ definido como

$$C = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{e}_i, 0 \leq \lambda_i \leq 1, \quad i = 1, 2, \ldots, n\}$$

se le llama *cubo unitario de n dimensiones* en $\mathbb{R}^n$. A los vectores $\mathbf{e}_i$ se les llama *lados del cubo* C, y a los vectores $\mathbf{d} \in \mathbb{R}^n$ de la forma $\mathbf{d} = \sum_{i=1}^{n} \alpha_i e_i$, donde $\alpha_i = \pm 1$, se les llama *diagonales del cubo* C (se identifica como la misma diagonal a los vectores $\mathbf{d}$ y $-\mathbf{d}$).

a. Dibuje un cubo unitario de una dimensión en $\mathbb{R}$.

b. Dibuje un cubo unitario de dos dimensiones en $\mathbb{R}^2$.

c. Dibuje un cubo unitario de tres dimensiones en $\mathbb{R}^3$.

d. Demuestre que el cubo unitario C de n dimensiones en $\mathbb{R}^n$ tiene 2^{n-1} diagonales distintas, las cuales tienen todas la misma longitud. ¿Cuál es esa longitud?

e. Sea $\mathbf{d}$ una diagonal del cubo C (en $\mathbb{R}^n$). Demuestre que si n es impar, no existen otras diagonales de C ortogonales a $\mathbf{d}$, en tanto que si n es par, digamos $n = 2k$, el cubo C tiene $\frac{1}{2}\binom{n}{k}$ diagonales ortogonales a $\mathbf{d}$.

f. Demuestre que el ángulo que forma una diagonal $\mathbf{d}$ del cubo C en $\mathbb{R}^n$ con cada uno de los lados del cubo es igual a $\arccos(n^{-1/2})$.

g. Demuestre que la norma de la proyección ortogonal de un lado cualquiera del cubo C en $\mathbb{R}^n$ sobre una diagonal cualquiera de C, es la n-ésima parte de la longitud de la diagonal $\mathbf{d}$.

(*) **40.** Considere las normas del máximo y de la suma

$$\|\mathbf{x}\|_{\text{máx}} = \text{máx}(|x_i|, i = 1, 2, \ldots, n) \qquad \|\mathbf{x}\|_s = |x_1| + |x_2| + \ldots + |x_n|$$

para un vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Demuestre que

$$\|\mathbf{x}\|_{\text{máx}} \leq \|\mathbf{x}\| \leq \|\mathbf{x}\|_s \leq n\|\mathbf{x}\|_{\text{máx}}$$

en donde $\|\mathbf{x}\|$ es la norma euclidiana de $\mathbf{x}$.

(*) **41.** Sea $\mathbf{x}_0 \in \mathbb{R}^n$ y $r > 0$. Se define la *bola abierta* (en $\mathbb{R}^n$) *con centro* en $\mathbf{x}_0$ y *radio* r, denotada por $B(\mathbf{x}_0, r)$, como el conjunto

$$B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| < r\}$$

a. ¿Cómo son las bolas abiertas en $\mathbb{R}$? Describa las bolas abiertas $B(2, 1)$ y $B(-3, 2)$.

b. ¿Cómo son las bolas abiertas en $\mathbb{R}^2$? Describa las bolas abiertas $B((2, 3), 1)$ y $B((-3, -1), 1)$.

c. ¿Cómo son las bolas abiertas en $\mathbb{R}^3$? Describa las bolas abiertas $B((0, 0, 0), 1)$ y $B((3, 5, 4), 2)$.

d. Suponga que en la definición dada de bola abierta tomamos la norma del máximo. Describa geométricamente la bola abierta en $\mathbb{R}^2$, $B((x_0, y_0), r)$.

e. Suponga que en la definición dada de bola abierta tomamos la norma de la suma. Describa geométricamente la bola abierta en $\mathbb{R}^3$, $B((x_0, y_0), r)$.

(*) **42.** Un conjunto $A \subset \mathbb{R}^n$ se dice ser *acotado* si existe un $c > 0$ tal que

$$\|\mathbf{x}\| < c \quad \forall \mathbf{x} \in A$$

a. Demuestre que el conjunto $A \subset \mathbb{R}^n$ es acotado si y sólo si existe una bola abierta en $\mathbb{R}^n$, $B(\mathbf{x}_0, r)$ tal que $A \subset B(\mathbf{x}_0, r)$.

b. Con la definición dada anteriormente, diremos que el conjunto A es acotado *según* la *norma euclidiana*. Si en tal definición tomamos lo norma del máximo o la norma de la suma, diremos que el conjunto es acotado *según* la *norma del máximo* o *de* la *suma*, respectivamente. Demuestre que son equivalentes:

1. el conjunto $A \subset \mathbb{R}^n$ es acotado según la norma euclidiana,
2. el conjunto $A \subset \mathbb{R}^n$ es acotado según la norma del máximo,
3. el conjunto $A \subset \mathbb{R}^n$ es acotado según la norma de la suma.

1.4 Bases ortonormales. Cambios de base

Diremos que un conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ en $\mathbb{R}^n$ es un conjunto *ortonormal* si es un conjunto ortogonal (i.e. los vectores son dos a dos ortogonales) y todos los vectores son unitarios (i.e. su norma es 1). Diremos que una base β de $\mathbb{R}^n$ es una *base ortonormal* si el conjunto de n vectores que la constituyen es un conjunto ortonormal. Así pues, la base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de $\mathbb{R}^n$ es una base ortonormal si

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

El ejemplo por excelencia de base ortonormal en $\mathbb{R}^n$ es la base canónica de este espacio, cuyo i-ésimo vector es $\mathbf{e}_i = (0, \ldots, 1, \ldots, 0)$ (el 1 en la i-ésima coordenada). Es claro que los vectores $\mathbf{e}_i$ son ortogonales dos a dos y que su norma es 1.

Dado un vector no nulo $\mathbf{v} \in \mathbb{R}^n$, diremos que el vector $\mathbf{u} = (\|\mathbf{v}\|)^{-1}\mathbf{v}$ es el vector $\mathbf{v}$ *normalizado*. Observe que este vector $\mathbf{u}$ es unitario, pues

$$\|\mathbf{u}\| = \left\|(\|\mathbf{v}\|)^{-1}\mathbf{v}\right\| = (\|\mathbf{v}\|)^{-1}\|\mathbf{v}\| = 1$$

Ejemplo 1. Consideremos el conjunto $\{\mathbf{v}_1, \mathbf{v}_2\}$ de vectores en $\mathbb{R}^2$ en donde $\mathbf{v}_1 = (-2, 4)$, $\mathbf{v}_2 = (6, 3)$. Este es un conjunto ortogonal (ver ejemplo 3 de la sección 2) y es por tanto una base de $\mathbb{R}^2$ (¿por qué?). Sin embargo, no es una base ortonormal, pues los vectores involucrados no son unitarios. Para tener una base ortonormal, podemos normalizar los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$. (Es claro que la propiedad de ortogonalidad entre ellos no se pierde con su normalización, ¿por qué?). Así, los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ en donde

$$\mathbf{u}_1 = (\|(-2, 4)\|)^{-1}(-2, 4) = (\sqrt{20})^{-1}(-2, 4) = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$
$$\mathbf{u}_2 = (\|(6, 3)\|)^{-1}(6, 3) = (\sqrt{45})^{-1}(6, 3) = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$$

constituyen una base ortonormal de $\mathbb{R}^2$. ■

Si $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de $\mathbb{R}^n$, entonces cada vector $\mathbf{v} \in \mathbb{R}^n$ se escribe de manera única como combinación lineal de los vectores de β, digamos $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. Si la base β es ortonormal, los escalares c_i se pueden calcular fácilmente: tomando producto punto del vector $\mathbf{v}$ con el i-ésimo vector de la base β nos queda

$$\mathbf{v} \cdot \mathbf{v}_i = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) \cdot \mathbf{v}_i = \sum_{j=1}^{n} c_j(\mathbf{v}_j \cdot \mathbf{v}_i) = c_i$$

Se tiene pues el siguiente resultado.

Teorema 1.4.1 Si $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base ortonormal del espacio $\mathbb{R}^n$, cada vector $\mathbf{v} \in \mathbb{R}^n$ se escribe en términos de esta base como

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{v} \cdot \mathbf{v}_2)\mathbf{v}_2 + \cdots + (\mathbf{v} \cdot \mathbf{v}_n)\mathbf{v}_n$$ ■

Ejemplo 2. En el ejemplo 1 se vio que la base $\beta = \{\mathbf{u}_1, \mathbf{u}_2\}$ donde

$$\mathbf{u}_1 = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right) \qquad \mathbf{u}_2 = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$$

es una base ortonormal de $\mathbb{R}^2$. Escribamos el vector $\mathbf{v} = (1, 7)$ en términos de esta base. Se tiene

$$\begin{aligned}\mathbf{v} &= (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2\\ &= \left(-\frac{1}{\sqrt{5}} + \frac{14}{\sqrt{5}}\right)\mathbf{u}_1 + \left(\frac{2}{\sqrt{5}} + \frac{7}{\sqrt{5}}\right)\mathbf{u}_2\\ &= \frac{13}{\sqrt{5}}\mathbf{u}_1 + \frac{9}{\sqrt{5}}\mathbf{u}_2\end{aligned}$$

■

Observe que si $\mathbf{u} \in \mathbb{R}^n$ es un vector unitario y $\mathbf{v}$ es un vector cualquiera de $\mathbb{R}^n$, la proyección de $\mathbf{v}$ sobre $\mathbf{u}$ es, según se discutió en la sección 2

$$\mathrm{PR}_{\mathbf{v}\to\mathbf{u}} = \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{u}\|^2}\mathbf{u} = (\mathbf{v}\cdot\mathbf{u})\mathbf{u}$$

de modo que si $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base ortonormal de $\mathbb{R}^n$, la expresión de un vector $\mathbf{v} \in \mathbb{R}^n$ en términos de esta base no es más que la suma de sus proyecciones sobre cada uno de los vectores de la base, como se ilustra en la figura siguiente

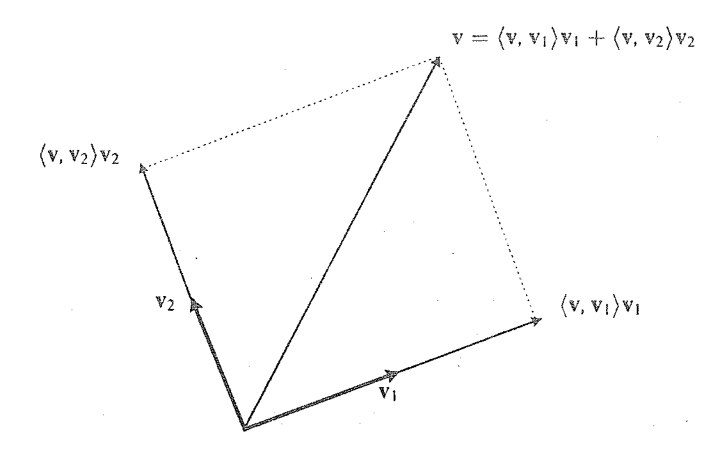


Figura 1. El vector $\mathbf{v}$ escrito en términos de la base ortonormal $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

Dada una base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ del espacio $\mathbb{R}^n$ es siempre posible, a partir de ella, construir una base ortonormal $\beta_{\text{on}} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ para este espacio. El proceso, conocido como *proceso* (de ortonormalización de bases) *de Gram-Schmidt,* se describe a continuación: como vector $\mathbf{u}_1$ tome al vector $\mathbf{v}_1$ normalizado. Es decir,

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1$$

Para construir el vector $\mathbf{u}_2$, consideramos la proyección del vector $\mathbf{v}_2$ sobre el vector unitario $\mathbf{u}_1$. Según lo discutido previamente esta proyección es $(\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$. Es claro entonces que el vector $\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$ es un vector ortogonal a $\mathbf{u}_1$. Normalizándolo, obtenemos el segundo vector de la base ortonormal procurada. Es decir

$$\mathbf{u}_2 = \frac{\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1}{\|\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1\|}$$

Podemos establecer este resultado considerando el vector $\mathbf{w} = \mathbf{v}_2 - a\mathbf{u}_1$, tomando a de tal modo que $\mathbf{w}$ sea ortogonal a $\mathbf{u}_1$, es decir, imponiendo la condición de que $(\mathbf{v}_2 - a\mathbf{u}_1) \cdot \mathbf{u}_1 = 0$. De aquí se obtiene que $a = \mathbf{v}_2 \cdot \mathbf{u}_1$, como antes. De manera análoga, para construír el vector $\mathbf{u}_3$, que debe ser ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$, consideramos el vector $\mathbf{w} = \mathbf{v}_3 - a\mathbf{u}_1 - b\mathbf{u}_2$, escogiendo a y b de modo que este vector sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$. Es decir, calculamos a y b imponiendo las condiciones siguientes:

$$(\mathbf{v}_3 - a\mathbf{u}_1 - b\mathbf{u}_2) \cdot \mathbf{u}_1 = 0$$
$$(\mathbf{v}_3 - a\mathbf{u}_1 - b\mathbf{u}_2) \cdot \mathbf{u}_2 = 0$$

De la primera de ellas se obtiene que $a = \mathbf{v}_3 \cdot \mathbf{u}_1$, y de la segunda $b = \mathbf{v}_3 \cdot \mathbf{u}_2$. Entonces, el vector $\mathbf{w} = \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2$ es ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$. Normalizándolo, obtenemos $\mathbf{u}_3$. Es decir

$$\mathbf{u}_3 = \frac{\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2}{\|\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2\|}$$

Continuando de esta manera, llegamos a que los vectores de la base ortonormal $\beta_{\text{on}} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, construida a partir de la base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ dada, son de la forma

$$\mathbf{u}_i = \frac{\mathbf{v}_i - (\mathbf{v}_i \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_i \cdot \mathbf{u}_2)\mathbf{u}_2 - \cdots - (\mathbf{v}_i \cdot \mathbf{u}_{i-1})\mathbf{u}_{i-1}}{\|\mathbf{v}_i - (\mathbf{v}_i \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_i \cdot \mathbf{u}_2)\mathbf{u}_2 - \cdots - (\mathbf{v}_i \cdot \mathbf{u}_{i-1})\mathbf{u}_{i-1}\|}$$

$i = 1, 2, \ldots, n$.

Ejemplo 3. Consideremos la base $\beta = \{(2, -1, 2), (3, 0, 4), (0, 1, 1)\}$ de $\mathbb{R}^3$. A partir de ella obtengamos una base ortonormal, usando el proceso de Gram-Schmidt presentado previamente. Llamando $\mathbf{v}_i$ a los vectores de la base β y $\mathbf{u}_i$, $i = 1, 2, 3$ a los de la base ortonormal procurada, tenemos que

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\|(2, -1, 2)\|}(2, -1, 2) = \frac{1}{3}(2, -1, 2)$$

Para determinar el vector $\mathbf{u}_2$, hagamos primero los cálculos siguientes

$$\begin{aligned}\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 &= (3, 0, 4) - \left((3, 0, 4) \cdot \frac{1}{3}(2, -1, 2)\right)\frac{1}{3}(2, -1, 2) \\ &= (3, 0, 4) - \frac{14}{9}(2, -1, 2) = \frac{1}{9}(-1, 14, 8)\end{aligned}$$

Normalizando este vector obtenemos $\mathbf{u}_2$

$$\mathbf{u}_2 = \frac{1}{\sqrt{261}}(-1, 14, 8)$$

Finalmente, para determinar $\mathbf{u}_3$ hacemos la siguiente operación

$$\begin{aligned} \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2 &= (0, 1, 1) - \left(\frac{1}{3}\right)\frac{1}{3}(2, -1, 2) - \left(\frac{22}{\sqrt{261}}\right)\frac{22}{\sqrt{261}}(-1, 14, 8) \\ &= (0, 1, 1) - \frac{1}{9}(2, -1, 2) - \frac{22}{261}(-1, 14, 8) = \frac{1}{29}(-4, -2, 3) \end{aligned}$$

Normalizando este vector, obtenemos $\mathbf{u}_3$

$$\mathbf{u}_3 = \frac{1}{\sqrt{29}}(-4, -2, 3)$$

Así, una base ortonormal de $\mathbb{R}^3$ es

$$\beta_{\text{on}} = \left\{\frac{1}{3}(2, -1, 2), \frac{1}{\sqrt{261}}(-1, 14, 8), \frac{1}{\sqrt{29}}(-4, -2, 3)\right\}$$

Si quisiéramos escribir el vector $\mathbf{v} = (3, 1, 5)$ como combinación lineal de los vectores de esta base, según lo establecido en el teorema 1.4.1, se tendría lo siguiente

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v} \cdot \mathbf{u}_3)\mathbf{u}_3 = 5\mathbf{u}_1 + \frac{51}{\sqrt{261}}\mathbf{u}_2 + \frac{1}{\sqrt{29}}\mathbf{u}_3$$ ■

Consideremos ahora el problema de cambio de bases ortonormales en el espacio $\mathbb{R}^n$. Por razones de simplicidad, vamos a estudiar de cerca solamente el caso $n = 3$. El caso general se copiará de éste "poniendo más coordenadas". Tomemos entonces dos bases ortonormales de $\mathbb{R}^3$, digamos $\beta_1 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\beta_2 = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$. Cada vector $\mathbf{v}$ en $\mathbb{R}^3$ se puede escribir como combinación de cada una de estas bases de la siguiente manera

$$\begin{aligned} \mathbf{v} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \\ \mathbf{v} &= d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + d_3\mathbf{u}_3 \end{aligned}$$

en donde sabemos que $c_i = \mathbf{v} \cdot \mathbf{v}_i$, $d_i = \mathbf{v} \cdot \mathbf{u}_i$, $i = 1, 2, 3$. Ciertamente también cada vector de la base β_2 se puede escribir en términos de los vectores de la base β_1, digamos que

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + a_{31}\mathbf{v}_3 \\ \mathbf{u}_2 &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + a_{32}\mathbf{v}_3 \\ \mathbf{u}_3 &= a_{13}\mathbf{v}_1 + a_{23}\mathbf{v}_2 + a_{33}\mathbf{v}_3 \end{aligned}$$

donde de hecho sabemos que $a_{ij} = \mathbf{u}_i \cdot \mathbf{v}_j$, $i, j = 1, 2, 3$. Escribiendo estas expresiones en la que expresa a $\mathbf{v}$ en términos de la base β_2 nos queda

$$\begin{aligned} \mathbf{v} &= d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + d_3\mathbf{u}_3 \\ &= d_1(a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + a_{31}\mathbf{v}_3) + d_2(a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + a_{32}\mathbf{v}_3) + d_3(a_{13}\mathbf{v}_1 + a_{23}\mathbf{v}_2 + a_{33}\mathbf{v}_3) \\ &= (a_{11}d_1 + a_{12}d_2 + a_{13}d_3)\mathbf{v}_1 + (a_{21}d_1 + a_{22}d_2 + a_{23}d_3)\mathbf{v}_2 + (a_{31}d_1 + a_{32}d_2 + a_{33}d_3)\mathbf{v}_3 \end{aligned}$$

Apoyados en la unicidad de la representación del vector v en términos de la base β_1, concluimos que

$$\begin{aligned} c_1 &= a_{11}d_1 + a_{12}d_2 + a_{13}d_3 \\ c_2 &= a_{21}d_1 + a_{22}d_2 + a_{23}d_3 \\ c_3 &= a_{31}d_1 + a_{32}d_2 + a_{33}d_3 \end{aligned}$$

Podemos escribir matricialmente estas expresiones como

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

A la matriz 3×1 de elementos c_i (d_i), se le llama "matriz de coordenadas del vector v respecto de la base β_1 (β_2 respectivamente)", la cual denotaremos por $[\mathbf{v}]_{\beta_1}$ ($[\mathbf{v}]_{\beta_2}$), y a la matriz 3×3 de elementos a_{ij}, en cuyas columnas aparecen los elementos de las matrices de coordenadas de cada vector de la base β_2 respecto de la base β_1, se le llama "matriz de cambio de base de β_2 a β_1". Denotaremos esta matriz por P. Se tiene entonces que

$$[\mathbf{v}]_{\beta_1} = P[\mathbf{v}]_{\beta_2}$$

Este es un resultado estándar en el problema de cambio de bases (no necesariamente ortonormales) en $\mathbb{R}^n$. Se puede ver también, por ejemplo, que la matriz P es inversible y que su inversa es la matriz de cambio de base de β_1 a β_2. Hasta este momento no hemos usado que nuestras bases son ortonormales. Este hecho se reflejará en las propiedades que en este caso tiene la matriz P de cambio de base. La característica fundamental de esta matriz en este caso es que su inversa coincide con su transpuesta. En efecto, hagamos el producto PP^t. Se tiene

$$PP^t = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

Llamemos δ_{ij} al elemento de la i-ésima línea y j-ésima columna de PP^t. Este es

$$\begin{aligned} \delta_{ij} &= a_{i1}a_{j1} + a_{i2}a_{j2} + a_{i3}a_{j3} \\ &= (\mathbf{u}_i \cdot \mathbf{v}_1)(\mathbf{u}_j \cdot \mathbf{v}_1) + (\mathbf{u}_i \cdot \mathbf{v}_2)(\mathbf{u}_j \cdot \mathbf{v}_2) + (\mathbf{u}_i \cdot \mathbf{v}_3)(\mathbf{u}_j \cdot \mathbf{v}_3) \\ &= \mathbf{u}_i \cdot ((\mathbf{u}_j \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}_j \cdot \mathbf{v}_2)\mathbf{v}_2 + (\mathbf{u}_j \cdot \mathbf{v}_3)\mathbf{v}_3) \\ &= \mathbf{u}_i \cdot \mathbf{u}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \end{aligned}$$

De modo entonces que PP^t es la matriz identidad, y así $P^{-1} = P^t$ como se quería ver. A una matriz inversible cuya inversa coincide con su transpuesta se le llama *matriz ortogonal*. Hemos entonces probado que la matriz de cambio de una base ortonormal a otra base ortonormal es una matriz ortogonal.

Ejemplo 4. Sea β_1 la base canónica de $\mathbb{R}^3$ y considere la base $\beta_2 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ de este espacio. Obtenemos la matriz P de cambio de base de β_2 a β_1 expresando cada vector de la base β_2 como combinación lineal de los vectores de la base β_1. Puesto que β_1 es la base canónica de

$\mathbb{R}^3$, las coordenadas de estas combinaciones lineales coinciden con las coordenadas de los vectores mismos (es decir, $[(x, y, z)]_{\beta_1} = (x, y, z)$). Entonces

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Por ejemplo, el vector $\mathbf{v} = (3, -1, 2) = 2(1, 1, 1) - 3(1, 1, 0) + 4(1, 0, 0)$, cuyas coordenadas en términos de la base β_2 son entonces $(2, -3, 4)$, se puede obtener como

$$P[(3, -1, 2)]_{\beta_2} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$$

Obsérvese que el resultado anterior nos da la versión del vector $\mathbf{v}$, cuyas coordenadas en la base β_2 son $(2, -3, 4)$, en términos de la base β_1. Como ésta es la base canónica, se obtiene directamente la expresión misma del vector $(3, -1, 2)$. La inversa de la matriz P,

$$P^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

es la matriz de cambio de base de β_1 a β_2. De este modo, el vector $(5, 1, 2)$, que en la base canónica tiene las mismas coordenadas, es transformado por la matriz P^{-1} a su "versión" en la base β_2. De hecho

$$P^{-1}[(5, 1, 2)]_{\beta_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$$

lo cual se comprueba fácilmente notando que $(5, 1, 2) = 2(1, 1, 1) - (1, 1, 0) + 4(1, 0, 0)$. ■

Ejemplo 5. En el espacio $\mathbb{R}^3$ considere la base canónica β_1 y la base ortonormal obtenida en el ejemplo 3

$$\beta_2 = \left\{ \frac{1}{3}(2, -1, 2), \frac{1}{\sqrt{261}}(-1, 14, 8), \frac{1}{\sqrt{29}}(-4, -2, 3) \right\}$$

La matriz de cambio de base de β_2 a β_1 es entonces

$$P = \begin{bmatrix} 2/3 & -1/\sqrt{261} & -4/\sqrt{29} \\ -1/3 & 14/\sqrt{261} & -2/\sqrt{29} \\ 2/3 & 8/\sqrt{261} & 3/\sqrt{29} \end{bmatrix}$$

Puesto que P es una matriz de cambio de una base ortonormal (β_2) a otra base ortonormal (la canónica), es una matriz ortogonal. Es decir, se debe tener que $P^{-1} = P^t$, como se puede comprobar fácilmente. ■

Ejemplo 6. Consideremos los vectores $\mathbf{e}_r = (\cos\theta, \operatorname{sen}\theta)$, $\mathbf{e}_\theta = (-\operatorname{sen}\theta, \cos\theta)$, donde $0 \leq \theta < 2\pi$. Puesto que

$$\mathbf{e}_r \cdot \mathbf{e}_\theta = (\cos\theta)(-\operatorname{sen}\theta) + (\operatorname{sen}\theta)(\cos\theta) = 0$$

$$\|\mathbf{e}_r\| = \|\mathbf{e}_\theta\| = \sqrt{\cos^2\theta + \operatorname{sen}^2\theta} = 1$$

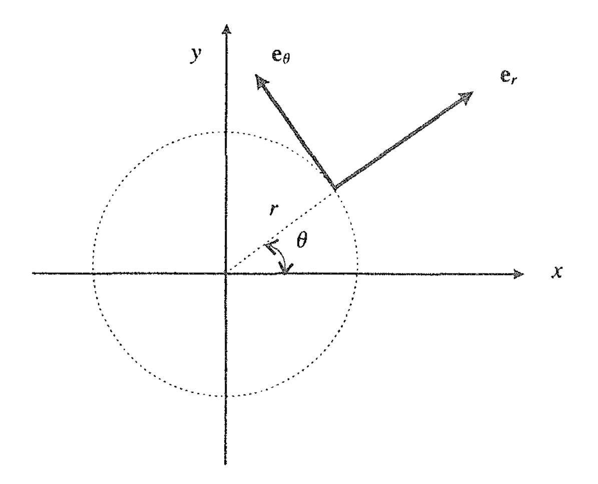


Figura 2. Los vectores unitarios $\mathbf{e}_r$ y $\mathbf{e}_\theta$.

la base $\beta_2 = \{\mathbf{e}_r, \mathbf{e}_\theta\}$ es una base ortonormal de $\mathbb{R}^2$. Esta es una base muy importante del espacio $\mathbb{R}^2$, pues cuando introducimos en él el sistema de coordenadas polares, los vectores $\mathbf{e}_r$ y $\mathbf{e}_\theta$ son los vectores unitarios que marcan las direcciones en que se miden las coordenadas r y θ en ese sistema. Sea $\beta_1 = \{i, j\}$ la base canónica de $\mathbb{R}^2$. Tenemos entonces que

$$\mathbf{e}_r = (\cos\theta, \operatorname{sen}\theta) = \cos\theta\mathbf{i} + \operatorname{sen}\theta\mathbf{j}$$
$$\mathbf{e}_\theta = (-\operatorname{sen}\theta, \cos\theta) = -\operatorname{sen}\theta\mathbf{i} + \cos\theta\mathbf{j}$$

y así, la matriz de cambio de base de β_2 a β_1 es

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

y, como sabemos, su transpuesta (su inversa) es la matriz de cambio de base de β_1 (la canónica) a β_2. Esta es

$$Q = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Tenemos pues que si $\mathbf{v} = (x, y) = x\mathbf{i} + y\mathbf{j}$ es un vector de $\mathbb{R}^2$, la multiplicación de Q por este vector nos da la expresión de $\mathbf{v}$ en términos de la base β_2, que denotamos como $[\mathbf{v}]_{\beta_2}$. Es decir

$$[\mathbf{v}]_{\beta_2} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta + y\operatorname{sen}\theta \\ -x\operatorname{sen}\theta + y\cos\theta \end{bmatrix}$$

Se tiene entonces que si $(x, y) = x\mathbf{i} + y\mathbf{j} \in \mathbb{R}^2$, entonces $(x, y) = (x\cos\theta + y\operatorname{sen}\theta)\mathbf{e}_r + (-x\operatorname{sen}\theta + y\cos\theta)\mathbf{e}_\theta$ en particular, poniendo $(x, y) = (1, 0) = \mathbf{i}$ y $(x, y) = (0, 1) = \mathbf{j}$, nos queda que

$$[\mathbf{i}]_{\beta_2} = \begin{bmatrix} \cos\theta \\ -\operatorname{sen}\theta \end{bmatrix} \qquad [\mathbf{j}]_{\beta_2} = \begin{bmatrix} \operatorname{sen}\theta \\ \cos\theta \end{bmatrix}$$

o sea

$$i = \cos\theta \mathbf{e}_r - \operatorname{sen}\theta \mathbf{e}_\theta$$
$$j = \operatorname{sen}\theta \mathbf{e}_r + \cos\theta \mathbf{e}_\theta$$

Por otra parte, podemos ver la base β_2 como la base canónica del sistema xy original girada un ángulo θ. Con esta idea se pueden obtener expresiones de un vector en el plano en términos de esta base (las "ecuaciones de transformación de coordenadas por rotación de ejes" que se estudian en los cursos de geometría analítica). ■

Ejercicios (Capítulo 1, Sección 4)

1. Determine el valor de a para que los siguientes conjuntos de vectores sean ortogonales.

   a. $\{(1, 2), (a, 5)\}$, **b.** $\{(0, 1), (1, a)\}$

   c. $\{(1, 2, -1), (3, 1, a)\}$, **d.** $\{(2, 3, 0), (0, 0, a)\}$

2. Verifique que el conjunto $\beta = \{(3/5, 4/5), (-4/5, 3/5)\}$ constituye una base ortonormal del espacio $\mathbb{R}^2$. Escriba el vector $\mathbf{v} = (6, -7)$ como combinación lineal de los elementos de esta base.

3. Compruebe que el conjunto

$$\beta = \{(2/3, -2/3), 1/3), (2/3, 1/3, -2/3), (1/3, 2/3, 2/3)\}$$

   constituye una base ortonormal del espacio $\mathbb{R}^3$. Escriba el vector $\mathbf{v} = (3, 1, 0)$ como combinación lineal de los elementos de esta base.

4. Verifique que el conjunto

$$\beta = \{(1/2, 1/2, 1/2, 1/2), (1/2, 1/2, -1/2, -1/2)$$
$$(1/2, -1/2, 1/2, -1/2), (1/2, -1/2, -1/2, 1/2)\}$$

   constituye una base ortonormal del espacio $\mathbb{R}^4$. Escriba el vector $\mathbf{v} = (2, 4, 1, 3)$ como combinación lineal de los elementos de esta base.

5. Considere la base canónica β' de $\mathbb{R}^2$. Determine la matriz de cambio de base de β a β', en donde β es la base ortonormal del ejercicio 2. Verifique que se trata de una matriz ortogonal. ¿Cuál es la matriz de cambio de base de β' a β?

6. Considere la base canónica β' de $\mathbb{R}^3$. Determine la matriz de cambio de base de β a β', en donde β es la base ortonormal del ejercicio 3. Verifique que se trata de una matriz ortogonal. ¿Cuál es la matriz de cambio de base de β' a β?

7. Considere la base canónica β' de $\mathbb{R}^4$. Determine la matriz de cambio de base de β a β', en donde β es la base ortonormal del ejercicio 4. Verifique que se trata de una matriz ortogonal. ¿Cuál es la matriz de cambio de base de β' a β?

En los ejercicios 8–11, aplique el proceso de Gram-Schmidt para obtener una base ortonormal a partir de la base dada del espacio $\mathbb{R}^2$.

8. $\beta = \{(2, 1), (1, 1)\}$

9. $\beta = \{(1, 1), (2, 4)\}$

10. $\beta = \{(2, 1), (1, -2)\}$

11. $\beta = \{(3, 1), (1, 1)\}$

En los ejercicios 12–15, siga el proceso de Gram-Schmidt para obtener una base ortonormal a partir de la base dada del espacio $\mathbb{R}^3$.

12. $\beta = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$

13. $\beta = \{(1, 2, 1), (3, 1, 2), (1, 0, 1)\}$

14. $\beta = \{(0, 1, 1), (1, 1, 0), (2, 1, 3)\}$

15. $\beta = \{(3, 1, 1), (-1, -1, -2), (1, 1, 4)\}$

En los ejercicios 16 y 17, use el proceso de Gram-Schmidt para obtener una base ortonormal a partir de la base dada del espacio $\mathbb{R}^4$.

16. $\beta = \{(1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0), (1, 0, 0, 0)\}$

17. $\beta = \{(1, 0, 1, -1), (2, 1, 3, 0), (0, 2, 1, 1), (0, 0, 1, 1)\}$

En los ejercicios 18–20, obtenga la matriz de cambio de base, de la base dada a la base canónica de $\mathbb{R}^3$. A partir de ella, obtenga la matriz de cambio de base de la canónica a la base dada.

18. $\beta = \{(2, 1, 1), (1, -1, 9), (1, 0, 0)\}$

19. $\beta = \{(1, 2, 3), (1, -1, -2), (3, 1, 1)\}$

20. $\beta = \{(0, 0, 5), (2, 3, 0), (1, 2, -1)\}$

21. Determine la matriz de cambio de base de la base ortonormal considerada en el ejercicio 2, a la base canónica de $\mathbb{R}^2$. Verifique que se trata de una matriz ortogonal.

22. Determine la matriz de cambio de base de la base ortonormal considerada en el ejercicio 3, a la base canónica de $\mathbb{R}^3$. Verifique que se trata de una matriz ortogonal.

23. Determine la matriz de cambio de base de la base ortonormal considerada en el ejercicio 4, a la base canónica de $\mathbb{R}^4$. Verifique que se trata de una matriz ortogonal.

1.5 El producto cruz en $\mathbb{R}^3$

En esta sección estudiaremos un nuevo producto entre vectores del espacio $\mathbb{R}^3$ (ya hemos considerado el producto punto entre estos vectores). Con él podremos estudiar las ecuaciones de planos en $\mathbb{R}^3$ que emprenderemos en la siguiente sección, junto con las ecuaciones de las rectas en $\mathbb{R}^3$. Una diferencia fundamental de este nuevo producto que estudiaremos ahora, es que éste será un *nuevo vector* de $\mathbb{R}^3$, mientras que el producto punto es, como sabemos, un escalar. Establezcamos la definición correspondiente.

Definición. Sean $\mathbf{u} = (x_1, x_2, x_3)$ y $\mathbf{v} = (y_1, y_2, y_3)$ dos vectores de $\mathbb{R}^3$. El *producto cruz* de $\mathbf{u}$ con $\mathbf{v}$, denotado por $\mathbf{u} \times \mathbf{v}$, es el vector de $\mathbb{R}^3$ dado por

$$\mathbf{u} \times \mathbf{v} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$$

o bien

$$\mathbf{u} \times \mathbf{v} = (x_2y_3 - x_3y_2)\mathbf{i} + (x_3y_1 - x_1y_3)\mathbf{j} + (x_1y_2 - x_2y_1)\mathbf{k}$$ ■

Si el lector siente que la fórmula de la definición anterior es demasiado para su memoria, no se preocupe. Lo es, de hecho, para la memoria de muchos de nosotros; además en matemáticas no se deben hacer esfuerzos mentales para memorizar fórmulas, sino para entender la esencia de las ideas, razonamientos y procedimientos analíticos que en ella aparecen. Con ayuda del siguiente determinante podremos recordar fácilmente la fórmula del producto cruz de dos vectores

$$\mathbf{u} \times \mathbf{v} = (x_1, x_2, x_3) \times (y_1, y_2, y_3) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$$

Desarrollando por cofactores a lo largo de la primera línea, la fórmula de la definición anterior se obtiene. Obsérvese que la matriz involucrada en este determinante es muy peculiar: en su primera linea están los vectores de la base canónica de $\mathbb{R}^3$, y las líneas restantes están ocupadas por escalares. Por supuesto que esto "no se vale" en matemáticas; sin embargo, el objetivo del determinante anterior es solamente proporcionar una ayuda para recordar la "difícil" fórmula del producto cruz de dos vectores. Además, con esta "manera de presentar la definición" se pueden probar fácilmente algunas de las propiedades del producto cruz que estudiaremos más adelante.

Una de las propiedades fundamentales que caracterizan al vector $\mathbf{u} \times \mathbf{v}$, es que éste es *perpendicular a ambos vectores* $\mathbf{u}$ y $\mathbf{v}$. En efecto, tomando el producto punto de $\mathbf{u} \times \mathbf{v}$ con $\mathbf{u}$ tenemos

$$\begin{aligned}(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} &= (x_2y_3 - x_3y_2)x_1 + (x_3y_1 - x_1y_3)x_2 + (x_1y_2 - x_2y_1)x_3 \\ &= x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_2x_1y_3 + x_3x_1y_2 - x_3x_2y_1 = 0\end{aligned}$$

Análogamente se ve que $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{v} = 0$.

Más aún, el vector $\mathbf{u} \times \mathbf{v}$ es un vector perpendicular a $\mathbf{u}$ y a $\mathbf{v}$, y su dirección se rige según la "regla de la mano derecha": usando la mano derecha, ponga su dedo pulgar perpendicular a los restantes dedos de la mano; haga coincidir la dirección de los cuatro dedos con la del vector $\mathbf{u}$ de tal modo que, cerrando su mano, ésta pueda encontrarse en su recorrido al vector $\mathbf{v}$. La dirección hacia donde apunta su dedo pulgar, es la del vector $\mathbf{u} \times \mathbf{v}$.

Ejemplo 1. Sean $\mathbf{u} = (2, 3, 8)$, $\mathbf{v} = (-2, 2, -5)$. Se tiene

$$\mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 3 & 8 \\ -2 & 2 & -5 \end{bmatrix} = -31\mathbf{i} - 6\mathbf{j} + 10\mathbf{k}$$

Además

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} = (-31, -6, 10) \cdot (2, 3, 8) = 0$$
$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{v} = (-31, -6, 10) \cdot (-2, 2, -5) = 0$$

como tenía que ocurrir. ■

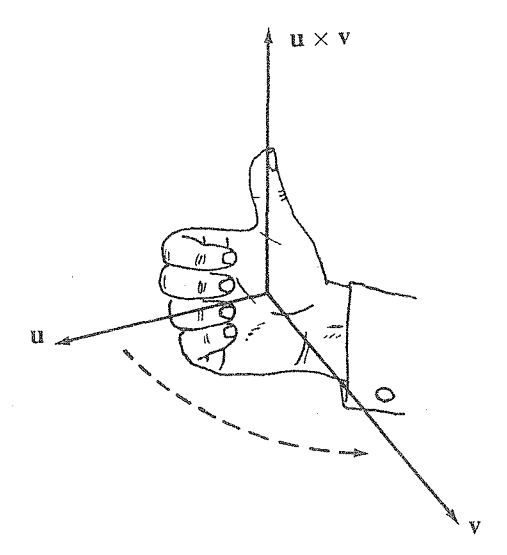


Figura 1. La regla de la mano derecha.

Con los vectores del ejemplo anterior, calculemos $\mathbf{v} \times \mathbf{u}$. Se tiene

$$\mathbf{v} \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -2 & 2 & -5 \\ 2 & 3 & 8 \end{bmatrix} = 31\mathbf{i} + 6\mathbf{j} - 10\mathbf{k}$$
$$= -(-31\mathbf{i} - 6\mathbf{j} + 10\mathbf{k}) = -\mathbf{u} \times \mathbf{v}$$

Así pues, el cambio de orden en los "factores" del producto cruz, produce un cambio de signo en el vector resultante. Este es un hecho general que se entiende fácilmente si se recuerda la propiedad de los determinantes de cambiar de signo cuando se intercambian de posición dos de sus líneas. Así, para dos vectores cualesquiera $\mathbf{u} = (x_1, x_2, x_3)$ y $\mathbf{v} = (y_1, y_2, y_3)$ se tiene

$$\mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} = -\det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ y_1 & y_2 & y_3 \\ x_1 & x_2 & x_3 \end{bmatrix} = -\mathbf{v} \times \mathbf{u}$$

Debido a esta propiedad, se dice que el producto cruz es *anticonmutativo*.

Al igual que el producto punto, el producto cruz tiene un comportamiento lineal respecto de sus dos variables (ver teorema 1.2.1, propiedades (3) y (4)). Con más presición, se tiene, con $\mathbf{u}$, $\mathbf{u}'$, $\mathbf{v}$, $\mathbf{v}' \in \mathbb{R}^3$ y $\lambda \in \mathbb{R}$

$$(\mathbf{u} + \lambda\mathbf{u}') \times \mathbf{v} = \mathbf{u} \times \mathbf{v} + (\lambda\mathbf{u}') \times \mathbf{v} = \mathbf{u} \times \mathbf{v} + \lambda(\mathbf{u}' \times \mathbf{v})$$
$$\mathbf{u} \times (\mathbf{v} + \lambda\mathbf{v}') = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times (\lambda\mathbf{v}') = \mathbf{u} \times \mathbf{v} + \lambda(\mathbf{u} \times \mathbf{v}')$$

Obsérvese entonces que el producto cruz, al igual que el producto punto, es *distributivo*. Nuevamente, esta propiedad se puede entender fácilmente recordando que la función determinante tiene un comportamiento lineal respecto de sus lineas (de hecho, también respecto de sus columnas). Es

decir, si $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$, $\mathbf{v}' = (y_1', y_2', y_3')$, se tiene

$$\begin{aligned}
\mathbf{u} \times (\mathbf{v} + \lambda \mathbf{v}') &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 + \lambda y_1' & y_2 + \lambda y_2' & y_3 + \lambda y_3' \end{bmatrix} \\
&= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} + \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ \lambda y_1' & \lambda y_2' & \lambda y_3' \end{bmatrix} \\
&= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} + \lambda \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1' & y_2' & y_3' \end{bmatrix} \\
&= \mathbf{u} \times \mathbf{v} + \lambda \mathbf{u} \times \mathbf{v}'
\end{aligned}$$

Como consecuencia de esta propiedad, podemos escribir, en general, que si $\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{v}_1, \ldots, \mathbf{v}_m$ son $(n + m)$ vectores de $\mathbb{R}^3$ y $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m$ son escalares, entonces

$$\left(\sum_{i=1}^{n} \lambda_i \mathbf{u}_i\right) \times \left(\sum_{j=1}^{n} \mu_1 \mathbf{v}_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \mu_j \mathbf{u}_i \times \mathbf{v}_j$$

Otra propiedad importante del producto cruz es su relación con la dependencia lineal de los vectores involucrados en él. Por una parte, es fácil verificar que si los vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ son linealmente dependientes (l.d.), entonces su producto cruz es cero (el vector cero de $\mathbb{R}^3$). En efecto, si $\mathbf{u}$ y $\mathbf{v}$ son l.d., existe un escalar $c \in \mathbb{R}$ tal que $\mathbf{u} = c\mathbf{v}$. Entonces, usando el comportamiento lineal del producto cruz previamente estudiado se tiene

$$\begin{aligned}
\mathbf{u} \times \mathbf{v} &= (c\mathbf{v}) \times \mathbf{v} = c(\mathbf{v} \times \mathbf{v}) \\
&= -\mathbf{v} \times \mathbf{u} = -\mathbf{v} \times (c\mathbf{v}) = -c(\mathbf{v} \times \mathbf{v})
\end{aligned}$$

o sea $2c(\mathbf{v} \times \mathbf{v}) = \mathbf{0}$, de donde $c = 0$ o bien $\mathbf{v} \times \mathbf{v} = \mathbf{0}$. De cualquier modo se tiene $\mathbf{u} \times \mathbf{v} = \mathbf{0}$. En particular, el producto cruz de un vector con él mismo es cero.

La afirmación recíproca de la verificada anteriormente también es cierta. Es decir, si $\mathbf{u}$ y $\mathbf{v}$ son vectores de $\mathbb{R}^3$ tales que $\mathbf{u} \times \mathbf{v} = \mathbf{0}$, entonces estos vectores son l.d. La validez de esta afirmación (y de la anterior también) se deduce fácilmente de la fórmula que a continuación obtendremos para la norma del producto cruz de dos vectores: si $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$ se tiene

$$\begin{aligned}
\|\mathbf{u} \times \mathbf{v}\|^2 &= (x_2 y_3 - x_3 y_2)^2 + (x_3 y_1 - x_1 y_3)^2 + (x_1 y_2 - x_2 y_1)^2 \\
&= x_2^2 y_3^2 + x_3^2 y_2^2 + x_3^2 y_1^2 + x_1^2 y_3^2 + x_1^2 y_2^2 + x_2^2 y_1^2 - 2x_2 x_3 y_2 y_3 - 2x_1 x_3 y_1 y_3 - 2x_1 x_2 y_1 y_2 \\
&= x_1^2 y_1^2 + x_1^2 y_2^2 + x_1^2 y_3^2 + x_2^2 y_1^2 + x_2^2 y_2^2 + x_2^2 y_3^2 + x_3^2 y_1^2 + x_3^2 y_2^2 + x_3^2 y_3^2 \\
&\quad - (x_1^2 y_1^2 + x_2^2 y_2^2 + x_3^2 y_3^2 + 2x_1 x_2 y_1 y_2 + 2x_1 x_3 y_1 y_3 + 2x_2 x_3 y_2 y_3) \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1 y_1 + x_2 y_2 + x_3 y_3)^2 \\
&= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \\
&= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \cos^2 \theta = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 (1 - \cos^2 \theta) \\
&= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \operatorname{sen}^2 \theta
\end{aligned}$$

o sea que

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| |\operatorname{sen} \theta|$$

donde se usó la fórmula $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$, siendo θ el ángulo que forman los vectores $\mathbf{u}$ y $\mathbf{v}$ (ver sección 3). Esta es una fórmula importante que nos descubrirá un hecho geométrico interesante sobre el significado de la norma del producto cruz de dos vectores. De ella se ve también que si $\mathbf{u}$ y $\mathbf{v}$ son vectores l.d., entonces $\operatorname{sen} \theta = 0$ (¿por qué?), y entonces $\|\mathbf{u} \times \mathbf{v}\| = \mathbf{0}$, de modo que $\mathbf{u} \times \mathbf{v} = \mathbf{0}$. Este hecho lo dedujimos anteriormente usando otras ideas. Por otro lado, si $\mathbf{u} \times \mathbf{v} = \mathbf{0}$, con $\mathbf{u}$ y $\mathbf{v}$ vectores no nulos de $\mathbb{R}^3$ (si alguno o ambos fueran nulos, su dependencia lineal se deduce de inmediato), entonces $\|\mathbf{u} \times \mathbf{v}\| = \mathbf{0}$, y así, por la fórmula anterior, $\operatorname{sen} \theta = 0$ (es decir, el ángulo entre $\mathbf{u}$ y $\mathbf{v}$ o es 0 o π; en cualquier caso, los vectores son colineales), y entonces $\mathbf{u}$ y $\mathbf{v}$ son l.d. como se quería ver.

En el siguiente teorema se recogen los resultados anteriormente estudiados acerca del producto cruz.

Teorema 1.5.1 Sean $\mathbf{u}$, $\mathbf{u}'$, $\mathbf{v}$, $\mathbf{v}'$ vectores de $\mathbb{R}^3$, y λ un escalar. Entonces

1. $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} = 0$
   $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{v} = 0$
2. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$ (anticonmutatividad)
3. $\mathbf{u} \times (\mathbf{v} + \lambda \mathbf{v}') = \mathbf{u} \times \mathbf{v} + \lambda \mathbf{u} \times \mathbf{v}'$
   $(\mathbf{u} + \lambda \mathbf{u}') \times \mathbf{v} = \mathbf{u} \times \mathbf{v} + \lambda \mathbf{u}' \times \mathbf{v}$ (distributividad)
4. $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ si y solamente si $\mathbf{u}$ y $\mathbf{v}$ son l.d.
5. $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| |\operatorname{sen} \theta|$, donde θ es el ángulo entre $\mathbf{u}$ y $\mathbf{v}$.

Demostración. Hecha en las líneas anteriores. Q.E.D.

Ejemplo 2. Consideremos los vectores $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ de la base canónica de $\mathbb{R}^3$. Es fácil verificar que $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$. Esquemáticamente podemos pensar en un triángulo como el siguiente

$$\begin{array}{ccc} & \mathbf{i} & \\ \swarrow & & \nwarrow \\ \mathbf{j} & \longrightarrow & \mathbf{k} \end{array}$$

El producto de cada uno de dos vértices (en el sentido indicado) es igual al tercer vértice. Por ejemplo, calculemos el producto cruz del vector $\mathbf{u} = (1, 0, -7)$ con $\mathbf{v} = (0, 2, -3)$. Podemos hacerlo de la manera "tradicional"

$$\mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & -7 \\ 0 & 2 & -3 \end{bmatrix} = 14\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$$

o bien, usando algunas propiedades del producto cruz de la manera siguiente

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= (\mathbf{i} - 7\mathbf{k}) \times (2\mathbf{j} - 3\mathbf{k}) = 2\mathbf{i} \times \mathbf{j} - 3\mathbf{i} \times \mathbf{k} - 14\mathbf{k} \times \mathbf{j} + 21\mathbf{k} \times \mathbf{k} \\ &= 2\mathbf{k} + 3\mathbf{k} \times \mathbf{i} + 14\mathbf{j} \times \mathbf{k} + 21(0) = 2\mathbf{k} + 3\mathbf{j} + 14\mathbf{i} \end{aligned}$$

■

La fórmula establecida en la propiedad 5) del teorema 1.5.1 anterior tiene una interpretación geométrica interesante: si consideramos el paralelogramo cuyos lados son los vectores $\mathbf{u}$ y $\mathbf{v}$, el

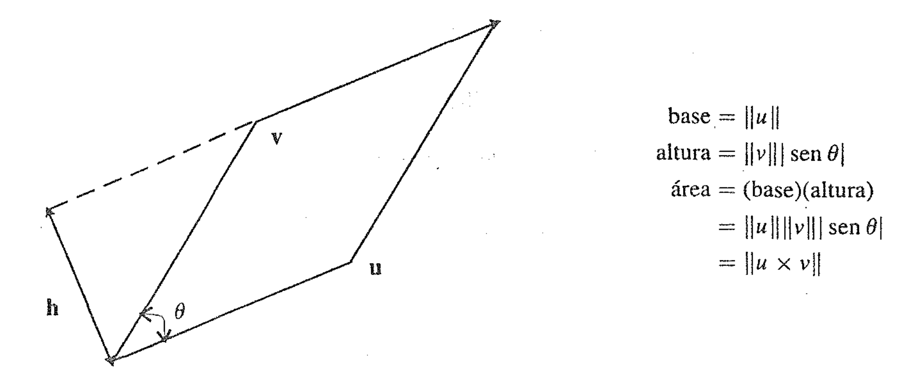


Figura 2. Paralelogramo generado por los vectores **u** y **v**.

área que éste encierra se calcula como la base (la norma de uno de los vectores) por la altura, la cual es la norma del otro vector multiplicada por el (valor absoluto del) seno del ángulo que forman los vectores (ver figura 2). Entonces, según la fórmula mencionada, el área de tal paralelogramo es justamente la norma del producto cruz de los vectores que lo generan.

Ejemplo 3. Consideremos los vectores $\mathbf{u} = (2, 3, 8)$, $\mathbf{v} = (-2, 2, -5)$ del ejemplo 1. Se tiene $\mathbf{u} \times \mathbf{v} = (-31, -6, 10)$, de modo que el área del paralelogramo generado por estos vectores es $\|\mathbf{u} \times \mathbf{v}\| = [(-31)^2 + (-6)^2 + (10)^2]^{1/2} = \sqrt{1097}$. Usando estas ideas se concluye inmediatamente que si los vectores $\mathbf{u}$ y $\mathbf{v}$ son vectores unitarios y ortogonales, entonces su producto cruz será también un vector unitario (y ortogonal a ambos), puesto que $\|\mathbf{u} \times \mathbf{v}\|$ es igual a $\|\mathbf{u}\|\|\mathbf{v}\| \,|\operatorname{sen}\theta| = (1)(1)\operatorname{sen}(\pi/2) = 1 =$ área de un cuadrado de lado 1. ∎

Combinemos ahora el producto cruz de dos vectores con el producto punto. Dados dos vectores $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, su producto cruz es un vector $\mathbf{v} \times \mathbf{w} \in \mathbb{R}^3$ con el que podemos tomar el producto punto con otro vector $\mathbf{u} \in \mathbb{R}^3$, para formar así el *escalar* $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$. A éste se le conoce como *triple producto* (escalar) de los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$. Si $\mathbf{u} = (x_1, x_2, x_3)$, $\mathbf{v} = (y_1, y_2, y_3)$, $\mathbf{w} = (z_1, z_2, z_3)$ y si $\mathbf{v} \times \mathbf{w} = (\alpha, \beta, \gamma)$, entonces $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = x_1\alpha + x_2\beta + x_3\gamma$. Recordando que α, β y γ eran los determinantes menores de la matriz cuya segunda y tercera líneas estaban formadas por las coordenadas de $\mathbf{v}$ y $\mathbf{w}$ asociados a los elementos de su primera línea, es claro que podemos escribir

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \det \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}$$

El escalar resultante del triple producto escalar de los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ tiene una interpretación geométrica interesante: consideremos el paralelepípedo generado por los tres vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ (cuyos ocho vértices son entonces (1) el origen, (2) el vector $\mathbf{u}$, (3) el vector $\mathbf{v}$, (4) el vector $\mathbf{w}$, (5) el vector $\mathbf{u} + \mathbf{v}$, (6) el vector $\mathbf{u} + \mathbf{w}$, (7) el vector $\mathbf{v} + \mathbf{w}$, (8) el vector $\mathbf{u} + \mathbf{v} + \mathbf{w}$), el cual tiene por volumen V al área A de su base multiplicada por su altura H. El área A de su base es el área del paralelogramo generado por los vectores $\mathbf{v}$ y $\mathbf{w}$ (véase figura 3). Esta es, según se vio anteriormente, la norma $\|\mathbf{v} \times \mathbf{w}\|$. La altura del paralelepípedo se puede calcular como la *norma de la proyección*

del vector restante **u** *sobre un vector perpendicular* a **v** y a **w**. Tal vector puede ser **v** × **w**, y la proyección de **u** sobre él es

$$\mathrm{PR}_{\mathbf{u}\to\mathbf{v}\times\mathbf{w}} = \frac{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})}{\|\mathbf{v}\times\mathbf{w}\|^2}\mathbf{v}\times\mathbf{w}$$

de modo que la altura H del paralelepípedo es

$$\begin{aligned} H = \|\,\mathrm{PR}_{\mathbf{u}\to\mathbf{v}\times\mathbf{w}}\,\| & \left|\frac{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})}{\|\mathbf{v}\times\mathbf{w}\|^2}\mathbf{v}\times\mathbf{w}\right| \\ &= \frac{|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|}{\|\mathbf{v}\times\mathbf{w}\|^2}\|\mathbf{v}\times\mathbf{w}\| \\ &= \frac{|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|}{\|\mathbf{v}\times\mathbf{w}\|} \end{aligned}$$

y entonces el volumen del paralelepípedo es

$$V = (\text{Área de la base})(\text{Altura}) = \|\mathbf{v}\times\mathbf{w}\|\frac{|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|}{\|\mathbf{v}\times\mathbf{w}\|} = |\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|$$

Es decir, el triple producto escalar **u** · (**v** × **w**) es, en valor absoluto, el volumen del paralelepípedo generado por los tres vectores **u**, **v** y **w**.

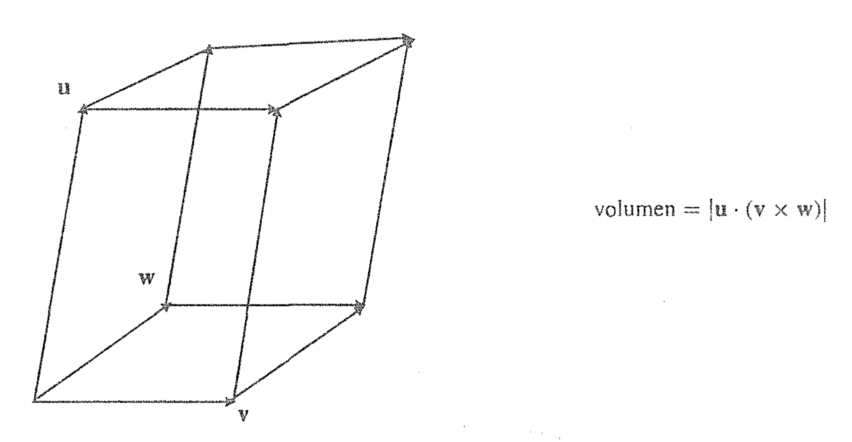


Figura 3. Paralelepípedo generado por los vectores **u**, **v** y **w**.

Ejemplo 4. Consideremos los vectores $\mathbf{u} = (2, 1, 0)$, $\mathbf{v} = (3, 4, 5)$, $\mathbf{w} = (2, 6, 4)$. El triple producto escalar de estos vectores es

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \det\begin{bmatrix} 2 & 1 & 0 \\ 3 & 4 & 5 \\ 2 & 6 & 4 \end{bmatrix} = 2(-14) - (2) = -30$$

Entonces el paralelepípedo generado por estos vectores tiene por volumen 30 unidades cúbicas. ■

En los ejercicios al final de esta sección se considerarán algunas propiedades de este triple producto escalar.

Apéndice Coordenadas cilíndricas y esféricas

En este apéndice introduciremos dos nuevos sistemas coordenados muy importantes en el espacio $\mathbb{R}^3$, los cuales jugarán un papel fundamental en algunos cálculos que se harán con integrales triples en el capítulo 6 (sección 7). De modo más concreto, los resultados que se establecerán en este apéndice, se usarán en el apéndice de la sección 2 del capítulo 7. En la discusión que presentamos aquí, usaremos libremente las ideas estudiadas en la sección anterior.

Dado un punto $\mathbf{p} = (x, y, z) \in \mathbb{R}^3$, asociamos a él la terna (r, θ, z) donde (r, θ) son las coordenadas polares de la proyección $\mathbf{p}'$ del punto $\mathbf{p}$ en el plano xy (es decir, (r, θ) son las coordenadas polares del punto $\mathbf{p}' = (x, y, 0)$). Decimos que los elementos de la terna (r, θ, z) son las *coordenadas cilíndricas* del punto $\mathbf{p}$ —ver figura 4—. (Obsérvese que, en cierto sentido, podemos decir que las coordenadas cilíndricas son las "coordenadas polares de $\mathbb{R}^3$", donde la tercera coordenada, que se encarga de medir la altura de un punto al plano xy, coincide con la del sistema rectangular).

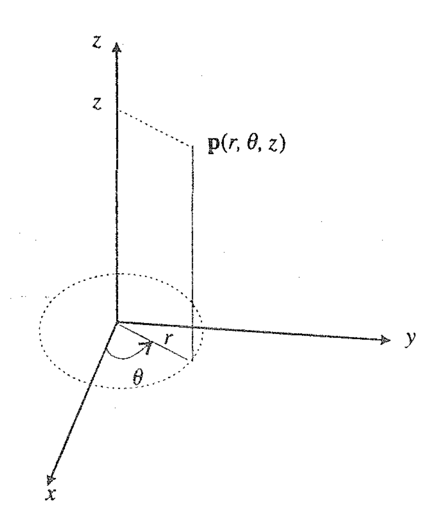


Figura 4. El sistema de coordenadas cilíndricas.

Obsérvese que la ecuación $r = r_0$ $(r_0 > 0)$ es, en el sistema cilíndrico, la ecuación de un *cilindro* (circular recto) cuyo eje es el eje z; en ella se establece que los puntos $\mathbf{p} = (r, \theta, z)$ que la satisfacen son puntos de $\mathbb{R}^3$ cuya proyección $\mathbf{p}'$ en el plano xy siempre dista del origen una cantidad constante igual a r_0. Si $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z$ son los vectores unitarios ortogonales que marcan la dirección donde se mide cada una de las correspondientes coordenadas r, θ, z, respectivamente en el sistema de coordenadas cilíndricas, (por lo cual $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ constituye una base ortonormal de $\mathbb{R}^3$), es claro que el vector $\mathbf{e}_z$ es justamente el vector $\mathbf{k} = (0, 0, 1)$ del sistema rectangular, y que los vectores $\mathbf{e}_r, \mathbf{e}_\theta$ son los mismos que los vectores del sistema ortonormal de $\mathbb{R}^2$ considerado en el ejemplo 4 de la sección anterior (con su última coordenada igual a cero). Es decir, se tiene que

$$\mathbf{e}_r = (\cos\theta, \operatorname{sen}\theta, 0) = \cos\theta\mathbf{i} + \operatorname{sen}\theta\mathbf{j}$$
$$\mathbf{e}_\theta = (-\operatorname{sen}\theta, \cos\theta, 0) = -\operatorname{sen}\theta\mathbf{i} + \cos\theta\mathbf{j}$$
$$\mathbf{e}_z = \mathbf{k}$$

Así pues, la matriz de cambio de base de la base ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$ a la base canónica de $\mathbb{R}^3$ es

$$P = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Siendo esta matriz ortogonal su transpuesta, que es justamente su inversa, nos da la matriz de cambio de base de la canónica a la base ortonormal del sistema de coordenadas cilíndricas. Esta es

$$P^{-1} = P^{\mathrm{t}} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta & 0 \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

de donde se obtiene que

$$\begin{aligned} \mathbf{i} &= \cos\theta \mathbf{e}_r - \operatorname{sen}\theta \mathbf{e}_\theta \\ \mathbf{j} &= \operatorname{sen}\theta \mathbf{e}_r + \cos\theta \mathbf{e}_\theta \\ \mathbf{k} &= \mathbf{e}_z \end{aligned}$$

Ejemplo 5. El punto $\mathbf{p}$ cuyas coordenadas cartesianas son $(1, 1, 1)$. Las coordenadas cilíndricas de este punto las determinamos convirtiendo a polares las coordenadas $(1, 1, 0)$ de la proyección de $\mathbf{p}$ sobre el plano xy. Como sabemos, se tiene $r = \sqrt{x^2 + y^2} = \sqrt{2}$, $\theta = \arctan \frac{y}{x} = \arctan 1 = \pi/4$. Así entonces, las coordenadas cilíndricas de $\mathbf{p}$ son $(\sqrt{2}, \pi/4, 1)$. Nótese que para ver qué coordenadas corresponden a $\mathbf{p}$ en el sistema ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$, podemos usar las fórmulas que relacionan a $\mathbf{i}, \mathbf{j}, \mathbf{k}$ con los vectores del sistema ortonormal de las coordenadas cilíndricas, como sigue

$$\begin{aligned} (1, 1, 1) &= \mathbf{i} + \mathbf{j} + \mathbf{k} \\ &= \cos\theta \mathbf{e}_r - \operatorname{sen}\theta \mathbf{e}_\theta + \operatorname{sen}\theta \mathbf{e}_r + \cos\theta \mathbf{e}_\theta + \mathbf{e}_z \\ &= (\cos\theta + \operatorname{sen}\theta)\mathbf{e}_r + (\cos\theta - \operatorname{sen}\theta)\mathbf{e}_\theta + \mathbf{e}_z \end{aligned}$$

Poniendo $\theta = \pi/4$ se obtiene

$$(1, 1, 1) = \sqrt{2}\mathbf{e}_r + \mathbf{e}_z$$

Es decir, en el sistema ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$, el punto $(1, 1, 1)$ se vería como $(\sqrt{2}, 0, 1)$, tal como la intuición geométrica nos dice que debería ocurrir. ■

Otro sistema coordenado muy importante en el espacio $\mathbb{R}^3$ es el que corresponde a las coordenadas esféricas. Este sistema localiza los puntos en el espacio tridimensional con los siguientes tres parámetros: la distancia del punto $\mathbf{p} = (x, y, z)$ al origen de coordenadas, denotada por r; el ángulo que forma la proyección del vector $\mathbf{p}$ en el plano xy, con la parte positiva del eje x, denotado por θ, y, el ángulo que forma el vector $\mathbf{p}$ con la parte positiva del eje z, denotado por ϕ. Se dice entonces que la terna (r, θ, ϕ) son las *coordenadas esféricas* del punto $\mathbf{p}$. Los rangos de variación de cada una de estas coordenadas será, atendiendo a la manera como se efectúa su medición: $0 \leq r < +\infty$, $0 \leq \theta < 2\pi$, $0 \leq \phi \leq \pi$.

Obsérvese que la ecuación $r = r_0$ es, en coordenadas esféricas, la ecuación de una esfera con centro en el origen y radio r_0. En efecto, a esta ecuación la satisfacen por los puntos en $\mathbb{R}^3$ cuya distancia al origen es igual a r_0, sin importar sus ángulos θ y ϕ, y estos puntos son justamente los de la esfera mencionada. También es fácil ver que la ecuación $\phi = \phi_0$, $(\phi_0 \neq \pi/2)$ es, en coordenadas

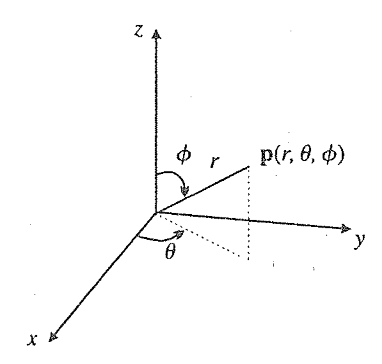


Figura 5. El sistema de coordenadas esféricas.

esféricas, la ecuación de un semicono con vértice en el origen, pues en ella se establece la condición de que los puntos **p** que la satisfacen forman un ángulo constante igual a ϕ_0 con la parte positiva del eje z, sin importar el ángulo θ y el valor de r.

Veamos la equivalencia entre las coordenadas rectangulares de un punto **p** en $\mathbb{R}^3$ y sus coordenadas esféricas. Si en el sistema rectangular **p** tiene coordenadas (x, y, z) y en el sistema esférico este punto tiene coordenadas (r, θ, ϕ), entonces, atendiendo a la figura 7 se tiene que

$$z = r\cos\phi$$

Del mismo modo, se tiene que

$$x = |\overline{\mathbf{Op}'}|\cos\theta, \qquad y = |\overline{\mathbf{Op}'}|\operatorname{sen}\theta$$

donde $\overline{\mathbf{Op}'} = r\operatorname{sen}\phi$. Entonces se tiene que $x = r\operatorname{sen}\phi\cos\theta$, $y = r\operatorname{sen}\phi\operatorname{sen}\theta$. Así pues, las fórmulas que relacionan las coordenadas (x, y, z) de un punto **p** en $\mathbb{R}^3$ con sus coordenadas (r, θ, ϕ)

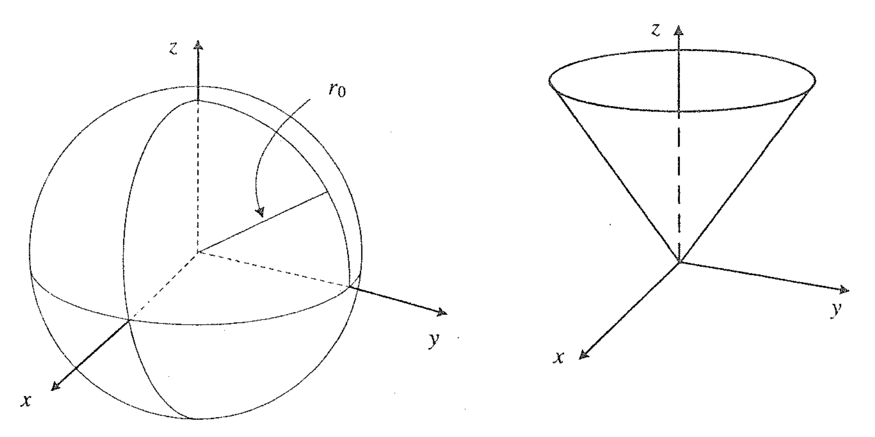


Figura 6. La esfera $r = r_0$ y el semicono $\phi = \phi_0$.

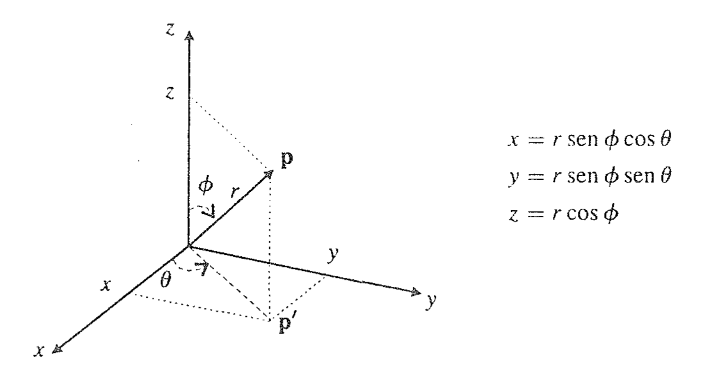


Figura 7. Relación entre las coordenadas cartesianas y esféricas.

son

$$x = r \operatorname{sen} \phi \cos \theta, \qquad y = r \operatorname{sen} \phi \operatorname{sen} \theta, \qquad z = r \cos \phi$$

Utilizando estas "fórmulas de transformación", se puede uno convencer de que la ecuación de una esfera $x^2 + y^2 + z^2 = r_0^2$, se ve, en el sistema esférico, como $r = r_0$. En efecto, al sustituir las expresiones anteriores en la ecuación de la esfera se obtiene $r^2 = r_0^2$, o bien $r = r_0$. Igualmente se ve que la ecuación del semicono $z = \sqrt{x^2 + y^2}$ se ve como $\phi = \pi/4$.

Obtengamos ahora las fórmulas que relacionan los vectores $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ del sistema ortonormal en las coordenadas cartesianas, con los vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\phi$ del sistema ortonormal esférico. Estos últimos marcan la dirección en que se efectúa la medición de coordenadas en el sistema esférico. Se tiene entonces que el vector $\mathbf{e}_r$ es un vector unitario en la dirección radial. Luego para $\mathbf{p} = (x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, se tiene

$$\mathbf{e}_r = \frac{1}{\|(x, y, z)\|}(x, y, z)$$

o bien, en términos de r, θ, ϕ se tiene

$$\begin{aligned} \mathbf{e}_r &= \frac{1}{r}(r \operatorname{sen} \phi \cos \theta, r \operatorname{sen} \phi \operatorname{sen} \theta, r \cos \phi) \\ &= (\operatorname{sen} \phi \cos \theta)\mathbf{i} + (\operatorname{sen} \phi \operatorname{sen} \theta)\mathbf{j} + (\cos \phi)\mathbf{k} \end{aligned}$$

El vector $\mathbf{e}_\theta$ es el mismo que el vector correspondiente (denotado igual) del ejemplo 4 (¿por qué?). Es decir,

$$\mathbf{e}_\theta = -\operatorname{sen} \theta\mathbf{i} + \cos \theta\mathbf{j}$$

Para determinar $\mathbf{e}_\phi$, obsérvese que se necesita un vector unitario que sea ortogonal a $\mathbf{e}_r$ y $\mathbf{e}_\theta$. Lo podemos obtener entonces como el producto cruz de $\mathbf{e}_\theta$ con $\mathbf{e}_r$. Puesto que estos dos últimos vectores

son unitarios, el vector $\mathbf{e}_\phi = \mathbf{e}_\theta \times \mathbf{e}_r$ también lo será (ver ejemplo 3). Así se tiene que

$$\mathbf{e}_\phi = \mathbf{e}_\theta \times \mathbf{e}_r = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ \operatorname{sen}\phi\cos\theta & \operatorname{sen}\phi\operatorname{sen}\theta & \cos\phi \end{bmatrix}$$
$$= (\cos\theta\cos\phi)\mathbf{i} + (\operatorname{sen}\theta\cos\phi)\mathbf{j} - (\operatorname{sen}\phi)\mathbf{k}$$

Entonces, las fórmulas que relacionan los vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\phi$ del sistema ortonormal en las coordenadas esféricas con los vectores $\mathbf{i}, \mathbf{j}, \mathbf{k}$ de la base canónica de $\mathbb{R}^3$ son

$$\begin{aligned} \mathbf{e}_r &= (\operatorname{sen}\phi\cos\theta)\mathbf{i} + (\operatorname{sen}\phi\operatorname{sen}\theta)\mathbf{j} + (\cos\phi)\mathbf{k} \\ \mathbf{e}_\theta &= -\operatorname{sen}\theta\mathbf{i} + \cos\theta\mathbf{j} \\ \mathbf{e}_\phi &= (\cos\theta\cos\phi)\mathbf{i} + (\operatorname{sen}\theta\cos\phi)\mathbf{j} - (\operatorname{sen}\phi)\mathbf{k} \end{aligned}$$

por lo que la matriz de cambio de base del sistema ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ a la base canónica de $\mathbb{R}^3$ es

$$P = \begin{bmatrix} \operatorname{sen}\phi\cos\theta & -\operatorname{sen}\theta & \cos\phi\cos\theta \\ \operatorname{sen}\phi\operatorname{sen}\theta & \cos\theta & \cos\phi\operatorname{sen}\theta \\ \cos\phi & 0 & -\operatorname{sen}\phi \end{bmatrix}$$

Su transpuesta (su inversa) es la matriz de cambio de base de la canónica a la base ortonormal del sistema de coordenadas esféricas

$$P^t = P^{-1} = \begin{bmatrix} \operatorname{sen}\phi\cos\theta & \operatorname{sen}\phi\operatorname{sen}\theta & \cos\phi \\ -\operatorname{sen}\theta & \cos\theta & 0 \\ \cos\phi\cos\theta & \cos\phi\operatorname{sen}\theta & -\operatorname{sen}\phi \end{bmatrix}$$

de donde se obtienen las relaciones

$$\begin{aligned} \mathbf{i} &= (\operatorname{sen}\phi\cos\theta)\mathbf{e}_r - (\operatorname{sen}\theta)\mathbf{e}_\theta + (\cos\phi\cos\theta)\mathbf{e}_\phi \\ \mathbf{j} &= (\operatorname{sen}\phi\operatorname{sen}\theta)\mathbf{e}_r + (\cos\theta)\mathbf{e}_\theta + (\cos\phi\operatorname{sen}\theta)\mathbf{e}_\phi \\ \mathbf{k} &= (\cos\phi)\mathbf{e}_r - (\operatorname{sen}\phi)\mathbf{e}_\phi \end{aligned}$$

Nótese, por último, que el sistema ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ está comprometido por la posición del punto $\mathbf{p} = (x, y, z)$, pues el vector $\mathbf{e}_r$ es un vector unitario en la *dirección de* $\mathbf{p}$. De hecho, el punto $\mathbf{p} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ se ve, en el sistema $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ como $\mathbf{p} = \sqrt{x^2 + y^2 + z^2}\mathbf{e}_r$, como el lector puede comprobar fácilmente.

Ejemplo 6. El punto $\mathbf{p} \in \mathbb{R}^3$ cuyas coordenadas esféricas son $(2, \pi/3, \pi/4)$ tiene, en el sistema coordenado rectangular, las coordenadas

$$\begin{aligned} x &= r\operatorname{sen}\phi\cos\theta = 2\operatorname{sen}\pi/4\cos\pi/3 = (2)(\sqrt{2}/2)(1/2) = \sqrt{2}/2 \\ y &= r\operatorname{sen}\phi\operatorname{sen}\theta = 2\operatorname{sen}\pi/4\operatorname{sen}\pi/3 = (2)(\sqrt{2}/2)(3/2) = \sqrt{6}/2 \\ z &= r\cos\phi = 2\cos\pi/4 = (2)(\sqrt{2}/2) = \sqrt{2} \end{aligned}$$

Es decir, en el sistema cartesiano el punto $\mathbf{p}$ se ve como $(\sqrt{2}/2, \sqrt{6}/2, \sqrt{2})$.

Por otra parte, teniendo las coordenadas rectangulares de un punto $\mathbf{p} \in \mathbb{R}^3$, podemos determinar sus coordenadas esféricas resolviendo para r, θ y ϕ las ecuaciones $x = r \operatorname{sen} \phi \cos \theta$, $y = r \operatorname{sen} \phi \operatorname{sen} \theta$, $z = r \cos \phi$: puesto que $x^2 + y^2 + z^2 = r^2$, se tiene $r = \sqrt{x^2 + y^2 + z^2}$; además, dividiendo y entre x se obtiene que $\tan \theta = y/x$, de donde $\theta = \arctan(y/x)$; por último, de la expresión $z = r \cos \phi$ obtenemos que $\phi = \arccos(z/r)$, donde $r = \sqrt{x^2 + y^2 + z^2}$. Por ejemplo, el punto $\mathbf{p}$ que en coordenadas rectangulares se ve como $(1, 2, -3)$, tiene, en el sistema esférico, las coordenadas

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} = \sqrt{14} \\ \theta &= \arctan(2/1) = \arctan 2 \\ \phi &= \arccos(-3/\sqrt{14}) \end{aligned}$$

■

Ejercicios (Capítulo 1, Sección 5)

En los ejercicios 1–5, calcule los productos cruz $\mathbf{u} \times \mathbf{v}$ y $\mathbf{v} \times \mathbf{u}$ de los vectores dados. En cada caso, verifique que el vector obtenido es ortogonal a cada uno de los vectores $\mathbf{u}$ y $\mathbf{v}$ dados.

1. $\mathbf{u} = (1, 1, 2)$, $\mathbf{v} = (-1, 1, 0)$
2. $\mathbf{u} = (2, 4, 3)$, $\mathbf{v} = (2, -4, 3)$
3. $\mathbf{u} = (0, 2, 5)$, $\mathbf{v} = (0, 4, 10)$
4. $\mathbf{u} = (2, 1, 1)$, $\mathbf{v} = (3, 2, 2)$
5. $\mathbf{u} = (3, 2, 2)$, $\mathbf{v} = (-3, -2, -2)$
6. Considere los vectores $\mathbf{u} = (3, 1, 2)$, $\mathbf{v} = (2, -4, 3)$, $\mathbf{w} = (1, 1, 7)$. Calcule los productos $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ y $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$. En base al resultado obtenido, explique por qué la expresión $\mathbf{u} \times \mathbf{v} \times \mathbf{w}$ para el producto cruz de tres vectores es una expresión ambigua.
7. Con los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ del ejercicio anterior, compruebe que $(\mathbf{u} + 2\mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + 2\mathbf{v} \times \mathbf{w}$.
8. ¿Verdadero o falso? Si $\mathbf{u} \times \mathbf{v} = \mathbf{u} \times \mathbf{w}$, entonces $\mathbf{v} = \mathbf{w}$.
9. Sean $\mathbf{u} = (2, -3, -3)$, $\mathbf{v} = (3, 1, 1)$. Calcule: a) $\mathbf{u} \times \mathbf{v}$; b) $(\mathbf{u} + \mathbf{v}) \times \mathbf{v}$; c) $(\mathbf{u} + \mathbf{v}) \times (\mathbf{u} + \mathbf{v})$; d) $(\mathbf{u} + \mathbf{v}) \times (\mathbf{u} - \mathbf{v})$; e) $(2\mathbf{u} + 3\mathbf{v}) \times (\mathbf{u} - 4\mathbf{v})$.
10. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ tres vectores tales que $\mathbf{u} + \mathbf{v} + \mathbf{w} = \mathbf{0}$. Demuestre que $\mathbf{u} \times \mathbf{v} = \mathbf{v} \times \mathbf{w} = \mathbf{w} \times \mathbf{u}$.
11. Sean $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4 \in \mathbb{R}^3$ cuatro vectores tales que $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{v}_3 \times \mathbf{v}_4$ y $\mathbf{v}_1 \times \mathbf{v}_3 = \mathbf{v}_2 \times \mathbf{v}_4$. Demuestre que los vectores $\mathbf{u} = \mathbf{v}_1 - \mathbf{v}_4$, $\mathbf{w} = \mathbf{v}_2 - \mathbf{v}_3$ son linealmente dependientes.
12. Demuestre que si los vectores $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ son colineales, entonces los vectores $\mathbf{u}$ y $\mathbf{v}$ son colineales. ¿Vale la afirmación recíproca?
13. Suponga que los vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ forman entre sí un ángulo de $\pi/4$. Demuestre que $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u} \times \mathbf{v}\|$.
14. Suponga que los vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ son vectores unitarios que forman entre sí un ángulo de $\pi/6$. Calcule $\|\mathbf{u} \times \mathbf{v}\|$.

15. Suponga que los vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ forman entre sí un ángulo de $\pi/6$. Si $\|\mathbf{u}\| = 6$, $\|\mathbf{v}\| = 5$, calcule $\|\mathbf{u} \times \mathbf{v}\|$.

16. Sean $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ dos vectores cuyas normas son 3 y 7 respectivamente. Si $\mathbf{u} \cdot \mathbf{v} = 5$, calcule $\|\mathbf{u} \times \mathbf{v}\|$.

17. Sean $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ dos vectores cuyas normas son 3 y 7 respectivamente. Si $\|\mathbf{u} \times \mathbf{v}\| = 5$, calcule $\mathbf{u} \cdot \mathbf{v}$.

18. Sean $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ dos vectores ortogonales con normas 4 y 2 respectivamente. Calcule $\|(\mathbf{u} + 2\mathbf{v}) \times (3\mathbf{u} - \mathbf{v})\|$.

19. Demuestre que para cualquiera de los dos vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ se cumple

$$\|\mathbf{u} \times \mathbf{v}\|^2 + (\mathbf{u} \cdot \mathbf{v})^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2.$$

(*) **20.** Sean $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ cuatro vectores en $\mathbb{R}^3$. Demuestre que

$$\mathbf{u} \cdot (\mathbf{v}_1 \times (\mathbf{v}_2 \times \mathbf{v}_3)) = (\mathbf{u} \times \mathbf{v}_1) \cdot (\mathbf{v}_2 \times \mathbf{v}_3)$$

(Sugerencia: use argumentos de desarrollo por cofactores a lo largo de las líneas de los determinantes con los que se calculan los triples productos escalares involucrados en esta fórmula).

b. Para los vectores $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$, demuestre que

$$(\mathbf{u} \times \mathbf{v}_1) \cdot (\mathbf{v}_2 \times \mathbf{v}_3) = \det \begin{bmatrix} \mathbf{u} \cdot \mathbf{v}_2 & \mathbf{u} \cdot \mathbf{v}_3 \\ \mathbf{v}_1 \cdot \mathbf{v}_2 & \mathbf{v}_1 \cdot \mathbf{v}_3 \end{bmatrix}$$

21. Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ n vectores en $\mathbb{R}^n$. Se define el *determinante de Gramm* de estos vectores, denotado por $\Gamma(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ como el determinante de la matriz $A = (a_{ij})_{i,j=1,\ldots,n}$, en donde $a_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j$. Observe que la matriz A es una matriz simétrica (es decir, A coincide con su transpuesta). Considere los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Demuestre que el cuadrado del triple producto escalar $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ es igual al determinante de Gramm de los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Es decir, demuestre que el volumen del paralelepípedo formado por los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$ es

$$|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| = \sqrt{\Gamma(\mathbf{u}, \mathbf{v}, \mathbf{w})}$$

(Sugerencia: considere la matriz A en cuyas líneas se encuentran las coordenadas de los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$; se sabe que $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| = |\det A|$; entonces $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|^2 = (\det A)^2 = (\det A)(\det A) = (\det A)(\det A^t) = \det AA^t$; concluya ...).

22. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tres vectores en $\mathbb{R}^3$. En este ejercicio probaremos la fórmula

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$$

a. Demuestre la fórmula anterior de manera directa, considerando las coordenadas de los vectores involucrados y efectuando las operaciones indicadas. Observe que es suficiente demostrar la fórmula con el vector $\mathbf{v}$ tomado sobre el eje x, ¿por qué?

b. Observe que el vector $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ es perpendicular a $\mathbf{v} \times \mathbf{w}$, quien a su vez es perpendicular a $\mathbf{v}$ y $\mathbf{w}$. Concluya entonces que el vector $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ es una combinación lineal $\mathbf{v}$ y $\mathbf{w}$. Es decir, concluya que existen constantes α y β tales que $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = \alpha\mathbf{v} + \beta\mathbf{w}$.

c. Tomando el producto punto del vector $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ con $\mathbf{u}$, concluya, de la expresión obtenida en el inciso anterior, que

$$(\mathbf{u} \cdot \mathbf{v})\alpha + (\mathbf{u} \cdot \mathbf{w})\beta = 0$$

d. Considere el vector $\mathbf{q} = (\alpha, \beta)$. Según el inciso anterior este vector es perpendicular al vector $\big((\mathbf{u} \cdot \mathbf{v}), (\mathbf{u} \cdot \mathbf{w})\big)$. Concluya entonces que el vector $\mathbf{p} = (-\beta, \alpha)$, el cual es perpendicular a $\mathbf{q}$, es paralelo a $((\mathbf{u} \cdot \mathbf{v}), (\mathbf{u} \cdot \mathbf{w}))$, y que por lo tanto existe una constante k tal que

$$(-\beta, \alpha) = k\big((\mathbf{u} \cdot \mathbf{v}), (\mathbf{u} \cdot \mathbf{w})\big)$$

e. Use el resultado del inciso anterior para reescribir la expresión obtenida en el inciso b) como

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = k((\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w})$$

f. Tome el producto punto en ambos miembros de la expresión obtenida en el inciso anterior, con un vector arbitrario $\omega \in \mathbb{R}^3$, y use los resultados del ejercicio 20, para demostrar que $k = 1$. Con esto se da un nuevo argumento que valida la fórmula establecida en este ejercicio.

23. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tres vectores en $\mathbb{R}^3$. Demuestre que

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{v} \cdot \mathbf{w})\mathbf{u}$$

24. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tres vectores en $\mathbb{R}^3$ tales que $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$. Discuta la posición relativa de los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$. (Sugerencia: use los resultados de los dos ejercicios anteriores).

25. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tres vectores linealmente independientes en $\mathbb{R}^3$. Demostrar que el plano que pasa por $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tiene por vector normal a

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} + \mathbf{v} \times \mathbf{w} + \mathbf{w} \times \mathbf{u}$$

26. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tres vectores en $\mathbb{R}^3$. Demuestre que

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0}$$

27. Demuestre que $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$, para cualesquiera vectores $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Explique entonces por qué la expresión $\mathbf{uvw}$ puede ser usada para denotar, sin peligro de confusión, al triple producto escalar de los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Téngase presente que usaremos esta notación en los ejercicios que siguen.

28. Demuestre que una condición necesaria y suficiente para que los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ sean coplanares (es decir, que se encuentren en un mismo plano) es que $\mathbf{uvw} = 0$.

En los ejercicios 29–32, determine si los vectores dados son coplanares o no. En caso de que lo sean, encuentre la ecuación del plano en que se encuentran.

29. $\mathbf{u} = (1, 2, 1)$, $\mathbf{v} = (1, -1, 0)$, $\mathbf{w} = (2, 1, 0)$

30. $\mathbf{u} = (2, 1, 1)$, $\mathbf{v} = (2, 3, 4)$, $\mathbf{w} = (2, -1, -2)$

31. $\mathbf{u} = (1, 1, 3)$, $\mathbf{v} = (3, 1, 1)$, $\mathbf{w} = (1, 3, 8)$

32. $\mathbf{u} = (2, 0, 1)$, $\mathbf{v} = (3, 1, 0)$, $\mathbf{w} = (2, 2, 3)$

En los ejercicios 33–35, demuestre que los cuatro puntos dados se encuentran en un mismo plano. Determine la ecuación del plano en que se encuentran.

33. $A = (1, 1, -1)$, $B = (0, 1, 1)$, $C = (1, 0, 1)$, $D = (2, 2, -5)$

34. $A = (-1, 1, 2)$, $B = (2, 2, 0)$, $C = (1, 1, 1)$, $D = (-1, 3, 1)$

35. $A = (0, 0, 1)$, $B = (2, -4, 3)$, $C = (5, -7, 2)$, $D = (-4, 7, -2)$

36. Sean $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tres vectores en $\mathbb{R}^3$. Demuestre que

$$|\mathbf{uvw}| \leq \|\mathbf{u}\|\|\mathbf{v}\|\|\mathbf{w}\|$$

37. Calcular el área del paralelogramo generado por los vectores $\mathbf{u} = (3, 2, 5)$, $\mathbf{v} = (1, 2, 7)$.

38. Calcular el área del paralelogramo cuyos vértices son $A = (1, 1, 1)$, $B = (2, 3, 4)$, $C = (-2, 1, 5)$, $D = (-1, 3, 8)$.

39. Calcular el área del triángulo cuyos vértices son $A = (3, 2, 3)$, $B = (-1, 2, 5)$, $C = (0, 2, 7)$.

40. Calcular el volumen del paralelepípedo generado por los vectores $\mathbf{u} = (2, 1, 4)$, $\mathbf{v} = (-1, 0, 9)$, $\mathbf{w} = (3, 2, 2)$.

41. Calcular el volumen del tetraedro cuyos vértices son el origen de coordenadas y los puntos $A = (2, 1, 1)$, $B = (-3, 7, 9)$ y $C = (-1, -5, 0)$.

42. Calcular el volumen del tetraedro cuyos vértices son los puntos $A = (2, 1, 2)$, $B = (5, 3, 7)$, $C = (-3, 4, 9)$ y $D = (10, 9, 11)$.

NOTA: *Los ejercicios 43–55 se refieren al material estudiado en el apéndice de esta sección (coordenadas cilíndricas y esféricas)*

43. Determine las coordenadas cilíndricas de los siguientes puntos dados en el sistema cartesiano: a. $\mathbf{p} = (2, 1, 1)$; b. $\mathbf{p} = (-1, 3, 5)$; c. $\mathbf{p} = (1, 0, 0)$; d. $\mathbf{p} = (2, 3, -1)$.

44. Determine las coordenadas cartesianas de los siguientes puntos dados en el sistema de coordenadas cilíndricas: a. $\mathbf{p} = (2, 0, 1)$; b. $\mathbf{p} = (1, \pi, 3)$; c. $\mathbf{p} = (3, 5\pi/3, -2)$; d. $\mathbf{p} = (16, \pi/4, 0)$.

45. Escriba la ecuación del cono $z = \sqrt{x^2 + y^2}$ en coordenadas cilíndricas.

46. Escriba la ecuación de la esfera unitaria $x^2 + y^2 + z^2 = 1$ en coordenadas cilíndricas.

47. Escriba la ecuación de los paraboloides: **a.** $z = x^2 + y^2$; **b.** $z = 2x^2 + 3y^2$, en coordenadas cilíndricas.

48. Determine las coordenadas esféricas de los siguientes puntos dados en el sistema coordenado cartesiano: **a.** $\mathbf{p} = (1, 0, 0)$; **b.** $\mathbf{p} = (3, 1, -1)$, **c.** $\mathbf{p} = (0, 1, 1)$; **d.** $\mathbf{p} = (-2, -3, -5)$.

49. Determine las coordenadas cartesianas de los siguientes puntos dados en el sistema de coordenads esféricas: **a.** $\mathbf{p} = (1, 0, 0)$; **b.** $\mathbf{p} = (2, \pi/2, \pi/2)$; **c.** $\mathbf{p} = (1, \pi/3, 3\pi/4)$; **d.** $\mathbf{p} = (4, 7\pi/4, \arccos(1/4))$.

50. Escriba la ecuación del cilindro circular recto $x^2 + y^2 = 9$ en coordenadas esféricas.

51. Escriba la ecuación de la esfera $(x - 1)^2 + y^2 + z^2 = 1$ en coordenadas esféricas.

52. Escriba la ecuación de las esferas: **a.** $(x - a)^2 + y^2 + z^2 = a^2$; **b.** $x^2 + (y - a)^2 + z^2 = a^2$; **c.** $x^2 + y^2 + (z - a)^2 = a^2$, en coordenadas esféricas.

53. Escriba la ecuación del cono $z^2 = a^2(x^2 + y^2)$ en coordenadas esféricas.

54. Sean P y Q dos puntos en $\mathbb{R}^3$, cuyas coordenadas en el sistema cilíndrico son (r_1, θ_1, z_1) y (r_2, θ_2, z_2), respectivamente. Demuestre que la distancia d entre estos dos puntos es

$$d = \sqrt{r_1^2 + r_2^2 - 2r_1r_2\cos(\theta_1 - \theta_2) + (z_2 - z_1)^2}$$

55. Sean P y Q dos puntos en $\mathbb{R}^3$, cuyas coordenadas en el sistema esférico son (r_1, ϕ_1, θ_1) y (r_2, ϕ_2, θ_2), respectivamente. Demuestre que la distancia d entre estos dos puntos es

$$d = \sqrt{r_1^2 + r_2^2 - 2r_1r_1[\operatorname{sen}\phi_1 \operatorname{sen}\phi_2\cos(\theta_1 - \theta_2) + \cos\phi_1\cos\phi_2]}$$

1.6 Rectas y planos en $\mathbb{R}^3$

En esta sección utilizaremos la herramienta vectorial desarrollada en las secciones anteriores (sobre todo las ideas en torno al del concepto de ortogonalidad) para hacer un poco de geometría analítica de planos y rectas en el espacio $\mathbb{R}^3$. Las generalizaciones al espacio $\mathbb{R}^n$ ($n > 3$) de los resultados que obtendremos serán casi inmediatas. La intención de esta sección es familiarizarnos con algunos tipos importantes de ecuaciones de planos y rectas en el espacio $\mathbb{R}^n$, pues todos los conceptos diferenciables del cálculo que se estudiarán en este libro tendrán su contenido geométrico al involucrar este tipo de lugares geométricos de $\mathbb{R}^n$, los que son, sin duda alguna, los más simples que existen. En el caso del plano $\mathbb{R}^2$, el lugar geométrico más simple que existe es la recta (en este caso no hay planos). Este es, por así decirlo, el prototipo de comportamiento "simple" y "decente" propio de una curva en el plano. La matemática involucrada en el estudio de estos "entes lineales" del plano es sencilla y cristalina; todos los resultados son "fáciles" de establecer alrededor de comportamientos lineales. Precisamente la intención del cálculo (diferencial) es *cambiar (localmente)* el *estudio de comportamientos "complicados" de funciones por comportamientos lineales los cuales* se *tienen "bien estudiados"*. Así, a una función $y = f(x)$ ("comportamiento complicado de x con y"), se le asigna, en un punto, una *recta tangente* ("comportamiento lineal que aproxima a la función") cuyas propiedades nos darán información de la función misma (en torno al punto en cuestión). Por esta razón es fundamental estar familiarizado con las ecuaciones de las rectas en el plano para poder entender las ideas del cálculo de funciones de una variable. En esta misma línea de razonamiento el objetivo de esta sección es presentar un resumen de las principales ideas en torno a los comportamientos lineales de lugares geométricos en el espacio $\mathbb{R}^n$, los cuales, insistimos, serán las "aproximaciones decentes" que tendrán (localmente) las funciones que estudiaremos en este libro.

Comencemos por la idea más simple: un plano Π en $\mathbb{R}^3$ queda completamente determinado si se conocen:

(*) un punto $\mathbf{p} = (x_0, y_0, z_0)$ por el que pasa

(*) un vector normal a él, digamos $\mathbf{n} = (a, b, c)$

Usaremos la palabra "normal" como sinónimo de "ortogonal" o "perpendicular". En efecto, ayudados con la figura 1 de inmediato se ve que el punto $\mathbf{q} = (x, y, z)$ pertenecerá al plano Π si y sólo si el vector diferencia $\mathbf{q} - \mathbf{p}$ se encuentra "sobre el plano". Es decir si y sólo si $\mathbf{q} - \mathbf{p} = (x - x_0, y - y_0, z - z_0)$ es ortogonal al vector $\mathbf{n} = (a, b, c)$ (el cual es claro que debe ser un vector no nulo). Así pues, el plano Π queda determinado como el lugar geométrico de aquellos puntos (x, y, z) de $\mathbb{R}^3$ tales que $(x - x_0, y - y_0, z - z_0) \cdot (a, b, c) = 0$. O sea tales que

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

Esta es la primera ecuación importante de la sección y merece que la encerremos en un cuadro con título para recordarla (¡la usaremos mucho!).

(1) $$\boxed{a(x - x_0) + b(y - y_0) + c(z - z_0) = 0}$$

Ecuación de un plano en $\mathbb{R}^3$ que pasa por (x_0, y_0, z_0) y tiene a $\mathbf{n} = (a, b, c)$ como vector normal

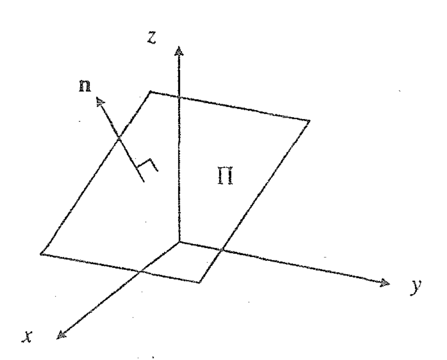


Figura 1. Un plano Π en $\mathbb{R}^3$ que pasa por $\mathbf{p}$ y tiene a $\mathbf{n}$ como vector normal.

La ecuación (1) se puede escribir, desarrollando los productos ahí indicados, como

$$ax + by + cz + d = 0$$

donde $d = -ax_0 - by_0 - cz_0$. Desde el punto de vista algebraico ésta es una ecuación *lineal* en las variables x, y, z. Si $d = 0$ es homogénea y, caso contrario, es no homogénea. Puesto que el vector $\mathbf{n} = (a, b, c)$ es no nulo, el hecho de que $d = 0$ significa que el plano Π *pasa por el origen*. Recuerde además que en el plano xy, las ecuaciones de las rectas eran justamente las ecuaciones lineales en las variables x, y (ecuaciones de la forma $ax + by = d$). Así pues, los planos en $\mathbb{R}^3$ son las "generalizaciones algebraicas naturales" de las rectas en el plano $\mathbb{R}^2$. Si vamos a dimensiones mayores, diremos que un *hiperplano* en el *espacio* $\mathbb{R}^n$ es el conjunto de puntos $(x_1, x_2, \ldots, x_n)$ en $\mathbb{R}^n$ que satisfacen la ecuación (lineal en las n variables $x_1, x_2, \ldots, x_n$)

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

donde el vector no nulo $\mathbf{n} = (a_1, a_2, \ldots, a_n)$ es un vector normal al hiperplano. Más aún, si $b = 0$, los vectores de este hiperplano constituyen un subespacio de $\mathbb{R}^n$ (ver ejemplo 1 de la sección 1) el cual tiene dimensión igual a $n - 1$ (¡verifique!).

Ejemplo 1. Hallemos la ecuación del plano en $\mathbb{R}^3$ que pasa por el punto $(2, 1, 3)$ y tiene a $\mathbf{n} = (2, -4, 5)$ por vector normal. Según (1) la ecuación procurada es $2(x-2)-4(y-1)+5(z-3) = 0$, o sea $2x - 4y + 5z = 15$. ■

Ejemplo 2. Consideremos el plano cuya ecuación es $7x - 2y + 6z = 12$. Es claro que el vector $\mathbf{n} = (7, -2, 6)$ es un vector normal al plano. Además, tal plano pasa, por ejemplo, por el punto $(0, 0, 2)$ (¿cómo determinar un punto por el que pasa el plano, conociendo su ecuación? Muy simple: fije los valores de dos de las tres variables involucradas en la ecuación y obtenga el valor de la variable restante; es obvio que los valores de x, y, z así obtenidos son las coordenadas de un punto que satisface la ecuación en cuestión). ■

Ejemplo 3. La ecuación $3x + 7y - 9z + 5u - 13w = 0$, representa un hiperplano en $\mathbb{R}^5$ que tiene a $\mathbf{n} = (3, 7, -9, 5, -13)$ por vector normal y pasa por el origen. Este es entonces un subespacio de $\mathbb{R}^5$ cuya dimensión es 4. Una base de él es, por ejemplo,

$$\beta = \{(-7, 3, 0, 0, 0), (3, 0, 1, 0, 0), (-5, 0, 0, 3, 0), (13, 0, 0, 0, 3)\}$$ ■

Supongamos ahora que nos dan tres puntos por los que pasa el plano Π, digamos $\mathbf{p}_1 = (x_1, y_1, z_1)$, $\mathbf{p}_2 = (x_2, y_2, z_2)$, $\mathbf{p}_3 = (x_3, y_3, z_3)$. Estos tres puntos pueden dejar bien determinado de la siguiente manera el plano Π que pasa por ellos. Si consideramos los vectores $\mathbf{u} = \mathbf{p}_2 - \mathbf{p}_1$ y $\mathbf{v} = \mathbf{p}_3 - \mathbf{p}_1$, éstos deben ser vectores que se encuentran *sobre el plano procurado*. Haciendo el producto cruz $\mathbf{u} \times \mathbf{v}$, obtenemos, en principio, un vector $\mathbf{n}$ normal al plano, a menos que este producto cruz sea cero (y en tal caso, los tres puntos dados no dejan bien determinado un plano que pase por ellos; piense, por ejemplo, en que los puntos sean colineales). Supongamos que efectivamente $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ es no nulo. En tal caso, nuestro problema queda resuelto: el plano Π que pasa por $\mathbf{p}_1$ (digamos) y tiene a $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ por vector normal está constituido por los puntos $\mathbf{q} = (x, y, z)$ de modo que el vector $\mathbf{q} - \mathbf{p}_1 = (x - x_1, y - y_1, z - z_1)$ es ortogonal a $\mathbf{n}$. Es decir,

$$(\mathbf{q} - \mathbf{p}_1) \cdot ((\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_3 - \mathbf{p}_1)) = 0$$

Obsérvese que se trata de un triple producto escalar. Escribiéndolo explícitamente, la ecuación procurada del plano es entonces

$$\det \begin{bmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{bmatrix} = 0$$

Ejemplo 4. Obtengamos la ecuación del plano que pasa por los tres puntos $\mathbf{p}_1 = (1, 1, 2)$, $\mathbf{p}_2 = (3, 6, 5)$, $\mathbf{p}_3 = (-2, 4, -8)$. Según la fórmula que acabamos de obtener, la ecuación procurada es

$$0 = \det \begin{bmatrix} x - 1 & y - 1 & z - 2 \\ 3 - 1 & 6 - 1 & 5 - 2 \\ -2 - 1 & 4 - 1 & -8 - 2 \end{bmatrix} = \det \begin{bmatrix} x - 1 & y - 1 & z - 2 \\ 2 & 5 & 3 \\ -3 & 3 & -10 \end{bmatrix}$$

o sea

$$-59(x - 1) + 11(y - 1) + 21(z - 2) = 0$$

o bien

$$59x - 11y - 21z - 6 = 0$$ ■

Ejemplo 5. Es claro que si $\mathbf{n} = (a, b, c)$ es un vector normal a un plano Π, cualquier vector del tipo $k\mathbf{n} = (ka, kb, kc)$, con $k \in \mathbb{R}$, $k \neq 0$, es también un vector normal al plano. Los planos $ax+by+cz = d$, $a'x+b'y+c'z = d'$ serán *paralelos* si y sólo si su vector normal es "esencialmente el mismo". Más aún, si el vector $\mathbf{n}' = (a', b', c')$, normal al segundo plano, es un múltiplo escalar del vector $\mathbf{n} = (a, b, c)$, normal al primer plano. Es decir, si existe $k \in \mathbb{R}$, $k \neq 0$, tal que $\mathbf{n}' = k\mathbf{n}$. Por ejemplo, los planos $2x - 5y + 4z = 1$, $8x - 20y + 16z = 7$ son paralelos, pues el vector $\mathbf{n}' = (8, -20, 16)$ normal al segundo plano es igual a 4 veces el vector $\mathbf{n} = (2, -5, 4)$, normal al primero. ■

Estudiaremos ahora las ecuaciones de las rectas en $\mathbb{R}^3$. Para determinar una recta en el espacio es suficiente conocer

(∗) un punto $\mathbf{p} = (x_0, y_0, z_0)$ por el que pasa

(∗) un vector $\mathbf{v} = (a, b, c)$ paralelo a la recta

En efecto, si $\mathbf{v} = (a, b, c)$ es un vector paralelo a la recta ℓ (digamos que es el vector que "marca la dirección de ℓ"), y la recta pasa por el punto $\mathbf{p} = (x_0, y_0, z_0)$, entonces el punto $\mathbf{q} = (x, y, z)$ pertenecerá a ℓ si y solamente si el vector $\mathbf{q} - \mathbf{p} = (x - x_0, y - y_0, z - z_0)$ (que se encuentra sobre la recta) es *paralelo* al vector $\mathbf{v}$ (ver figura 2). Es decir, debe existir $t \in \mathbb{R}$ (para cada punto $\mathbf{q}$), tal que $\mathbf{q} - \mathbf{p} = t\mathbf{v}$. Haciendo explícita esta condición se tiene

$$(x - x_0, y - y_0, z - z_0) = t(a, b, c)$$

de donde $x = x_0 + at$, $y = y_0 + bt$, $z = z_0 + ct$, donde $t \in \mathbb{R}$ (mientras t recorre $\mathbb{R}$, el punto $\mathbf{q} = (x, y, z)$ recorre la recta). Estas ecuaciones se conocen como las *ecuaciones paramétricas de la recta* (el número real t es "el parámetro") y, por su importancia, las dejamos escritas dentro de un cuadro

$$\text{(2)} \qquad \boxed{\begin{cases} x = x_0 + ta \\ y = y_0 + tb \\ z = z_0 + tc \end{cases} \quad t \in \mathbb{R}}$$

Ecuaciones paramétricas de la recta que pasa por (x_0, y_0, z_0) y tiene a $\mathbf{v} = (a, b, c)$ por vector paralelo

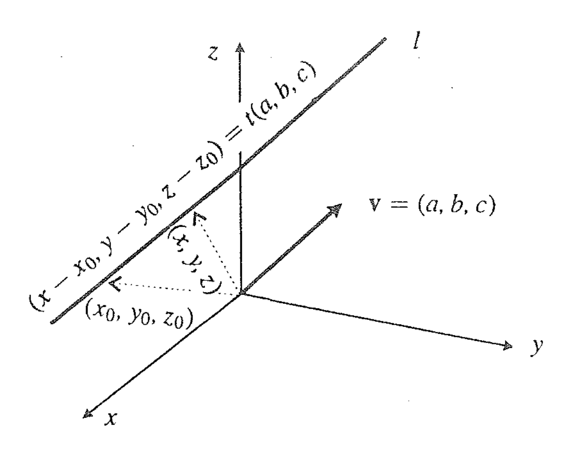


Figura 2. La recta ℓ que pasa por (x_0, y_0, z_0) y tiene a $\mathbf{v} = (a, b, c)$ por vector paralelo.

Nótese que, despejando e igualando el parámetro t de cada una de las ecuaciones paramétricas de la recta, la ecuación anterior también se puede escribir también como

$$\frac{x-x_0}{a}=\frac{y-y_0}{b}=\frac{z-z_0}{c}$$

Veamos algunos ejemplos.

Ejemplo 6. Obtengamos la ecuación de la recta que pasa por el punto $\mathbf{p}=(2,-3,7)$ y tiene a $\mathbf{v}=(-4,-6,1)$ por vector paralelo. Según la ecuación (2), la recta procurada es

$$\begin{cases} x=2-4t \\ y=-3-6t \qquad t\in\mathbb{R} \\ z=7+t \end{cases}$$ ■

Consideremos dos planos $a_1x+b_1y+c_1z=d_1$, $a_2x+b_2y+c_2z=d_2$ que no sean paralelos. La intersección de estos dos planos es una recta cuya ecuación se puede determinar resolviendo simultáneamente las ecuaciones de los planos. Estas son dos ecuaciones con tres indeterminadas, lo que permite dejar escritas dos de ellas en términos de la tercera. El ejemplo siguiente ilustra cómo proceder en este caso.

Ejemplo 7. Obtengamos la ecuación de la recta de la intersección de los dos planos $2x-3y+10z=1$, $x-y-z=2$. Estos planos no son paralelos (¿por qué?). Podemos, por ejemplo, dejar la variable z como "variable libre", digamos $z=t\in\mathbb{R}$, y resolver el sistema

$$2x-3y=1-10t, \qquad x-y=2+t$$

para x y y. Obtenemos que $x=5+13t$, $y=3+12t$. Así pues, la recta buscada tiene por ecuaciones paramétricas a

$$\begin{cases} x=5+13t \\ y=3+12t \qquad t\in\mathbb{R} \\ z=t \end{cases}$$

y se trata de una recta que pasa por el punto $\mathbf{p}=(5,3,0)$ y tiene a $\mathbf{v}=(13,12,1)$ por vector paralelo. ■

Ejemplo 8. Consideremos las dos rectas

$$\ell_1=\begin{cases} x=1+2t \\ y=-2-3t \\ z=5+4t \end{cases} \qquad \ell_2=\begin{cases} x=7+3s \\ y=2+2s \\ z=1-2s \end{cases}$$

con $t,s\in\mathbb{R}$ (por razones que en este ejemplo se harán evidentes, conviene denotar de manera distinta a los parámetros de las rectas). Nos preguntamos si existe algún punto común a ambas rectas. Es decir, si las rectas se intersectan en algún punto. Para que esto ocurriera, tendría que existir un valor de t y uno de s que hicieran que la x, y, z de la primera recta fuera igual a las correspondientes de la segunda. Es decir, que se cumpliera

$$\begin{aligned} 1+2t&=7+3s \\ -2-3t&=2+2s \\ 5+4t&=1-2s \end{aligned}$$

En este caso se ve fácilmente (¡haciendo las cuentas!) que este sistema de tres ecuaciones con dos indeterminadas tiene solución para $t = 0$ y $s = -2$. Sustituyendo estos valores de los parámetros en las ecuaciones de las rectas, encontramos que el punto común de ambas es $(1, -2, 5)$. Más aún, sabiendo ya que ambas rectas se intersectan, nos preguntamos por el plano que las contiene, es decir, el plano en el que se hallan simultáneamente ambas rectas. Para lograrlo, necesitamos un punto por el que pasa el plano (lo cual no es problema, pues *cualquier* punto de *cualquiera* de las dos rectas es un punto del plano) y un vector normal a él. Es claro que este vector debe ser ortogonal a todos los vectores que se encuentran *sobre* las rectas. En particular, debe ser perpendicular a los dos vectores paralelos a ambas rectas, que son $\mathbf{v}_1 = (2, -3, 4)$ y $\mathbf{v}_2 = (3, 2, -2)$. El producto cruz de estos vectores nos sirve entonces como vector normal al plano que buscamos. Este es

$$\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2 = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & -3 & 4 \\ 3 & 2 & -2 \end{bmatrix} = -2\mathbf{i} + 16\mathbf{j} + 13\mathbf{j}$$

De modo pues que el plano que contiene a ambas rectas es (tomando, por ejemplo, el punto $(1, -2, 5)$ que es común a ambas rectas como punto por el que pasa el plano)

$$-2(x - 1) + 16(y + 2) + 13(z - 5) = 0$$

o sea

$$-2x + 16y + 13z - 31 = 0$$

Llamamos la atención al hecho de que hubiéramos podido escoger como punto por el que pasa el plano, cualquier punto de cualquiera de las dos rectas y el resultado debe ser el mismo. Obsérvese que si tomamos un punto arbitrario de la primera recta, o sea $(1 + 2t, -2 - 3t, 5 + 4t)$, el plano que pasa por este punto y tiene a $\mathbf{v} = (-2, 16, 13)$ por vector normal es

$$-2(x - (1 + 2t)) + 16(y - (-2 - 3t)) + 13(z - (5 + 4t)) = 0$$

o sea, simplificando $-2x + 16y + 13z - 31 = 0$. Así mismo ocurre si tomamos cualquier punto de la segunda recta. ■

Ejemplo 9. Consideremos las dos rectas

$$\ell_1 = \begin{cases} x = -7 + 3t \\ y = -4 + 4t \\ z = -3 - 2 \end{cases} \qquad \ell_2 = \begin{cases} x = 21 + 6s \\ y = -5 - 4s \\ z = 2 - s \end{cases}$$

Al estudiar el sistema

$$\begin{aligned} -7 + 3t &= 21 + 6s \\ -4 + 4t &= -5 - 4s \\ -3 - 2t &= 2 - s \end{aligned}$$

vemos que éste no tiene solución. Esto significa que tales rectas no se intersectan. Además, es fácil convencerse de que tales rectas no son paralelas, pues sus vectores paralelos $\mathbf{v}_1 = (3, 4, -2)$ y $\mathbf{v}_2 = (6, -4, -1)$ son linealmente independientes. Nos planteamos el problema de encontrar la distancia más corta que existe entre estas dos rectas. Este es un problema interesante que puede ser atacado desde puntos de vista muy diversos con distintas herramientas. Por ejemplo, se puede resolver

(usando el cálculo que estudiaremos a partir del próximo capítulo) como un problema de extremos de funciones de dos variables (ver sección 6 del capítulo 4, ejercicios 43–46). Ahora presentamos un argumento "sin cálculo" que resuelve el problema. La idea general es "meter" la recta ℓ_1 en un plano Π de modo que la recta ℓ_2 quede paralela a este plano; logrando esto, la distancia procurada no es más que la distancia de un punto (cualquiera) de la recta ℓ_2 al plano Π. Los detalles son los siguientes: los vectores $\mathbf{v}_1 = (3, 4, -2)$ y $\mathbf{v}_2 = (6, -4, -1)$ son paralelos a las rectas ℓ_1 y ℓ_2 respectivamente. Procuramos un vector (a, b, c) que sea ortogonal a ambos. Se debe cumplir entonces que

$$\begin{aligned} 3a + 4b - 2c &= 0 \\ 6a - 4b - c &= 0 \end{aligned}$$

(este sistema tiene una infinidad de soluciones, lo cual, desde el punto de vista geométrico resulta perfectamente explicable), de donde, por ejemplo el vector $\mathbf{n} = (4, 3, 12)$ es un vector como el que procuramos. El plano que contiene a ℓ_1 y tiene a $\mathbf{n}$ por vector normal es $4x + 3y + 12z + 76 = 0$. Es claro que ℓ_2 quedó paralela a este plano. Tomando entonces un punto cualquiera de ℓ_2, digamos $(21, -5, 2)$, y calculando la distancia de éste al plano obtenido (la fórmula de la distancia de un punto a un plano se obtiene en el ejercicio 34 al final de esta sección) se llega a

$$d = \frac{|4(21) + 3(-5) + 12(2)|}{\sqrt{(4)^2 + (3)^2 + (12)^2}} = 13$$

Esta es la distancia procurada.

Con las mismas ideas manejadas en este ejemplo, se puede demostrar que la distancia d entre dos rectas no paralelas ℓ_1 y ℓ_2 viene dada por

$$d = \frac{|(\mathbf{p}_2 - \mathbf{p}_1) \cdot (\mathbf{v}_1 \times \mathbf{v}_2)|}{\|\mathbf{v}_1 \times \mathbf{v}_2\|}$$

en donde $\mathbf{p}_1$ y $\mathbf{p}_2$ son dos puntos cualesquiera sobre las rectas ℓ_1 y ℓ_2, respectivamente, y $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores paralelos a estas rectas. Dejamos al lector como ejercicio que dé el argumento que valide esta fórmula. ■

Ejercicios (Capítulo 1, Sección 6)

En los ejercicios 1–5, determine la ecuación del plano que pasa por el punto $\mathbf{p}$ y tiene al vector $\mathbf{n}$ como vector normal.

1. $\mathbf{p} = (0, 0, 0)$, $\mathbf{n} = (1, 1, 1)$

2. $\mathbf{p} = (2, 1, 1)$, $\mathbf{n} = (1, 0, 0)$

3. $\mathbf{p} = (3, 4, 5)$, $\mathbf{n} = (0, 2, 3)$

4. $\mathbf{p} = (2, -1, 0)$, $\mathbf{n} = (3, 2, 6)$

5. $\mathbf{p} = (0, 2, 0)$, $\mathbf{n} = (-2, -7, 4)$

6. Hallar la ecuación del plano que pasa por $\mathbf{p} = (x_0, y_0, z_0)$ y tiene a $\mathbf{p}$ por vector normal.

7. Considere los puntos $\mathbf{p} = (1, -1, 3)$, $\mathbf{q} = (3, 2, 1)$. Hallar la ecuación del plano: **a.** que pasa por $\mathbf{p}$ y tiene a $\mathbf{n} = \mathbf{p} - \mathbf{q}$ por vector normal; **b.** que pasa por $\mathbf{q}$ y tiene a $\mathbf{n} = \mathbf{q} - \mathbf{p}$ por vector normal.

8. Hallar la ecuación del plano que pasa por el punto $\mathbf{p} = (5, 1, 1)$, si se sabe que los vectores $\mathbf{u} = (2, 1, 2)$, $\mathbf{v} = (-4, -5, 7)$ son paralelos a él.

9. Hallar la ecuación del plano que pasa por los dos puntos $\mathbf{p} = (1, 1, 0)$, $\mathbf{q} = (3, 2, 4)$, si se sabe que el vector $\mathbf{u} = (7, -1, -3)$ es paralelo a él.

En los ejercicios 10–14, determine si los puntos $\mathbf{p}$ y $\mathbf{q}$ pertenecen al plano dado.

10. $3x - y + z = 1$, $\mathbf{p} = (0, 0, 1)$, $\mathbf{q} = (1, 1, -1)$

11. $z = 3$, $\mathbf{p} = (3, 1, 3)$, $\mathbf{q} = (3, 3, 5)$

12. $x + y - 4z = 0$, $\mathbf{p} = (0, 0, 0)$, $\mathbf{q} = (2, 2, 1)$

13. $3x - 2y = 0$, $\mathbf{p} = (2, 1, 1)$, $\mathbf{q} = (-3, 2, 5)$

14. $x + y - 2z = 10$, $\mathbf{p} = (5, 7, 2)$, $\mathbf{q} = (5, 7, 1)$

En los ejercicios 15–18, determine un punto por el que pasa el plano dado y un vector normal a él.

15. $3x + z = 3$

16. $y = 0$

17. $x - y - z = 5$

18. $3x - 2y + 7z = 23$

En los ejercicios 19–23, determine si los planos dados son paralelos, perpendiculares, o si no están en ninguno de estos dos casos. (Nota: dos planos son perpendiculares si sus vectores normales lo son).

19. $3x + y - z = 3$, $z - y = 8$

20. $x + 4y - 2z = 1$, $2x + 8y - 4z = 7$

21. $y = 3$, $y = 7$

22. $x = 0$, $z = 0$

23. $x - y + z = 1$, $x - y + z = 9$

24. Hallar la ecuación del plano que pasa por el punto $\mathbf{p} = (3, 2, 2)$ y es paralelo al plano $3x - 2y + z = 6$.

25. Hallar la ecuación del plano que pasa por el origen de coordenadas y es perpendicular al plano $4x - y + z = 9$.

En los ejercicios 26–30, determine la ecuación del plano que pasa por los tres puntos dados.

26. $\mathbf{p} = (0, 0, 0)$, $\mathbf{q} = (3, 1, 1)$, $\mathbf{r} = (-1, 2, 4)$

27. $\mathbf{p} = (2, 1, 0)$, $\mathbf{q} = (0, 0, 7)$, $\mathbf{r} = (2, 1, 1)$

28. $\mathbf{p} = (1, -1, -1)$, $\mathbf{q} = (8, 4, 2)$, $\mathbf{r} = (2, 1, 5)$

29. $\mathbf{p} = (1, 4, 9)$, $\mathbf{q} = (-3, 1, 5)$, $\mathbf{r} = (4, 4, 11)$

30. $\mathbf{p} = (a, 0, 0)$, $\mathbf{q} = (0, b, 0)$, $\mathbf{r} = (0, 0, c)$

31. Demuestre que la ecuación de un plano que pasa por el punto $\mathbf{p} = (x_0, y_0, z_0)$, tal que los vectores $\mathbf{u} = (a_1, b_1, c_1)$, $\mathbf{v} = (a_2, b_2, c_2)$ (no colineales) son paralelos a él, se puede escribir como

$$\det \begin{bmatrix} x + x_0 & y - y_0 & z - z_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix} = 0$$

32. Demuestre que la ecuación de un plano que pasa por los puntos $\mathbf{p} = (x_0, y_0, z_0)$ y $\mathbf{q} = (x_1, y_1, z_1)$, tal que el vector $\mathbf{u} = (a, b, c)$ es un vector paralelo a él, se pude escribir como

$$\det \begin{bmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ a & b & c \end{bmatrix} = 0$$

33. Considere el plano que pasa por el origen $Ax + By + Cz = 0$, y sea $\mathbf{p} = (x_0, y_0, z_0)$ un punto que no pertenece al plano. Use la proyección del vector $\mathbf{p}$ sobre el vector (ortogonal al plano) $\mathbf{n} = (A, B, C)$, para demostrar que la distancia perpendicular del punto $\mathbf{p}$ al plano es

$$d = \frac{|Ax_0 + By_0 + Cz_0|}{\sqrt{A^2 + B^2 + C^2}}$$

34. Demuestre que la distancia (perpendicular) del punto $\mathbf{p} = (x_0, y_0, z_0)$ al plano $Ax + By + Cz + D = 0$, es

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

En los ejercicios 35–37, calcule la distancia del punto $\mathbf{p}$ al plano dado.

35. $\mathbf{p} = (5, 30, 426)$, $x = 3$

36. $\mathbf{p} = (3, -2, 5)$, $2x - y + z = 0$

37. $\mathbf{p} = (1, 1, 5)$, $2x + 3y - 2z = 4$

38. Habiendo verificado que los planos $2x + y - z = 4$, $4x + 2y - 2z - 5 = 0$ son paralelos, calcule la distancia entre ellos.

39. Suponga que los planos $A_1x + B_1y + C_1z = D_1$, $A_2x + B_2y + C_2z = D_2$ son paralelos. Obtenga una fórmula para calcular la distancia entre ellos.

40. Dos caras de un cubo se encuentran en los planos $3x - y + 2z = 5$, $3x - y + 2z = 7$. Calcule el volumen del cubo.

41. Demuestre que los planos paralelos al plano $Ax + By + Cz = D$ que distan de éste r unidades, son $Ax + By + Cz = D \pm r\sqrt{A^2 + B^2 + C^2}$.

42. Suponga que los planos perpendiculares $x + y - 2z = 2$, $2x + z = 5$ dividen a un cubo de volumen 64 en cuatro paralelepípedos. Si el centro del cubo se encuentra en el punto $(2, 2, 1)$, determine las ecuaciones de los planos en donde se encuentran las caras del cubo.

43. Los vectores $\mathbf{u} = (1, 2, 1)$, $\mathbf{v} = (-3, 1, 1)$, $\mathbf{w} = (1, -4, 7)$ determinan 3 de las aristas de un paralelepípedo. Halle las ecuaciones de los planos en que se encuentran sus caras.

44. Hallar un punto en el eje x que equidiste de los dos planos paralelos $3x - y + 2z = 6$, $3x - y + 2z = 13$.

45. Hallar un punto en el eje y que equidiste de los dos planos $2x + 2y + z = 0$, $4x - 3y = 2$.

46. Demuestre que los tres planos $x + y + z = 6$, $x - y - z = 0$, $2x - 3y + z = 1$ se cortan en un solo punto. Determine este punto.

47. Determine los puntos donde se encuentran los vértices del cubo del ejercicio 42.

48. Determine los puntos en que se encuentran los vértices del paralelepípedo del ejercicio 43.

En cada uno de los ejercicios 49–51, determine la ecuación de la recta que pasa por el punto $\mathbf{p}$ dado y tiene al vector $\mathbf{v}$ como vector paralelo.

49. $\mathbf{p} = (0, 0, 0)$, $\mathbf{v} = (1, 1, 1)$

50. $\mathbf{p} = (0, 1, 0)$, $\mathbf{v} = (0, 1, 0)$

51. $\mathbf{p} = (2, -4, -7)$, $\mathbf{v} = (3, 1, 2)$

52. Hallar la ecuación de la recta que pasa por el punto $\mathbf{p} = (2, 1, 1)$ y es paralela al vector que une $\mathbf{p}$ con el punto $\mathbf{q} = (2, -3, -5)$.

53. Hallar la ecuación de la recta que pasa por el punto $\mathbf{p} = (3, 4, 7)$ y es perpendicular al plano $3x - 2y + z = 9$.

54. Hallar la ecuación de la recta que pasa por el punto $\mathbf{p} = (x_0, y_0, z_0)$ y tiene a $\mathbf{p}$ por vector paralelo.

En los ejercicios 55–57, determine la ecuación de la recta que pasa por los dos puntos dados.

55. $\mathbf{p} = (3, 9, 7)$, $\mathbf{q} = (-1, 2, 5)$

56. $\mathbf{p} = (2, 1, 6)$, $\mathbf{q} = (-2, 3, 2)$

57. $\mathbf{p} = (0, 0, 0)$, $\mathbf{q} = (2, 6, 5)$

58. Determine las ecuaciones de las rectas donde se encuentran las diagonales del cubo del ejercicio 42. Halle el punto donde éstas se cruzan.

59. Determine las ecuaciones de las rectas en quee se encuentran las diagonales del paralelepípedo del ejercicio 43. Halle el punto donde éstas se cruzan.

En los ejercicios 60–62, determine si los puntos $\mathbf{p}$ y $\mathbf{q}$ se encuentran en la recta dada.

60. $\begin{cases} x = 2 + t \\ y = -3t \\ z = t \end{cases}$ $\quad \mathbf{p} = (2, 0, 0)$, $\mathbf{q} = (3, 1, 1)$

61. $\begin{cases} x = 1 + 2t \\ y = 2 - 3t \\ z = 1 + t \end{cases}$ $\quad \mathbf{p} = (5, -4, 1)$, $\mathbf{q} = (-1, 5, 0)$

63. $\begin{cases} x = 2 + t \\ y = -3t \\ z = t \end{cases}$ $\mathbf{p} = (0, 0, 0)$, $\mathbf{q} = (2, 1, 1)$

63. Hallar la ecuación de la recta que pasa por $\mathbf{p} = (2, 1, 4)$ y que es paralela a la recta $x = 3t$, $y = -2 + 4t$, $z = -t$.

64. Hallar la ecuación de la recta que pasa por el origen y es perpendicular a la recta $x = 3 - 2t$, $y = 3 + 4t$, $z = -5t$.

65. Los puntos $A = (2, 1, 3)$, $B = (-2, 7, 5)$, $C = (2, 3, 2)$ son los vértices de un triángulo. Hallar las ecuaciones de las rectas donde se encuentran las medianas de este triángulo (es decir, las rectas que salen de uno de los vértices hacia el punto medio del lado opuesto de él). Constate que estas tres rectas se cruzan en un punto.

66. Hallar los puntos de intersección de la recta $x = 3 + t$, $y = 2 - t$, $z = 4 - 5t$, con los planos coordenados.

67. Hallar el punto de intersección de la recta $\frac{x-3}{2} = \frac{y-1}{3} = z$, con el plano $2x + y - z = 1$.

68. Verifique que la recta $\frac{x-2}{2} = \frac{y+1}{7} = \frac{z}{4}$ se encuentra contenida en el plano $x - 2y + 3z - 4 = 0$.

69. Compruebe que la recta $\frac{x}{-5} = \frac{y+1}{12} = \frac{z-1}{13}$ se encuentra contenida tanto en el plano $5x + y + z = 0$, como en el plano $2x + 3y - 2z = -5$.

En cada uno de los ejercicios 70–73, determine las ecuaciones paramétricas de las rectas que resultan de la intersección de los planos dados.

70. $2x + 3y - z - 4 = 0$, $3x + y - z = 0$.

71. $3x + y - 4z = 0$, $5x + z = 2$.

72. $x + y + z = 2$, $x - y + z = 3$.

73. $x = 0$, $y = 0$.

74. Verifique que las dos rectas $L_1 = \{x = 3t, y = 2t, z = t, t \in \mathbb{R}\}$, $L_2 = \{x = -3t, y = -t, z = t, t \in \mathbb{R}\}$ se cortan en un punto. Determine la ecuación del plano en el que éstas se encuentran.

75. El punto $\mathbf{p} = (2, 1, -1)$ se encuentra en el plano $x - y + z = 0$. Determine la forma general de las ecuaciones de las rectas que pasan por $\mathbf{p}$ y que se encuentran sobre el plano dado.

76. El punto $\mathbf{p} = (1, 3, 2)$ se encuentra en el plano $x + y - 2z = 0$. Determine la forma general de las ecuaciones de las rectas que pasan por $\mathbf{p}$ y que se encuentran sobre el plano dado.

77. Hallar la distancia entre los puntos de intersección de la recta $x = 3 - 2t$, $y = z = t$, con los planos paralelos $2x + y + z = 3$, $2x + y + z = 9$. ¿Es ésta la distancia entre los dos planos paralelos dados?

78. Hallar la distancia entre los puntos de intersección de la recta $x = 5 + t$, $y = 3 - 2t$, $z = 4 + 3t$, con los planos paralelos $x - 2y + 3z = 2$, $x - 2y + 3z = 6$. ¿Es ésta la distancia entre los dos planos paralelos dados?

En los ejercicios 79–82 compruebe que las rectas dadas son paralelas.

79. $\begin{cases} x = 5 + 3t \\ y = -2 + 5t \\ z = t \end{cases}$, $\begin{cases} x = 7 - 6t \\ y = -10t \\ z = 4 - 2t \end{cases}$

80. $\begin{cases} x = 3 + t \\ y = 4t \\ z = 2 - 2t \end{cases}$, $\begin{cases} 2x + z = 5 \\ 4x + y + 4z = 7 \end{cases}$

81. $\begin{cases} x = 1 - t \\ y = 3 + 2t \\ z = -2 + 5t \end{cases}$, $\begin{cases} -x - 3y + z = 6 \\ 2x - 4y + 2z = 3 \end{cases}$

82. $\begin{cases} 2x + y - z = 7 \\ x + y = 6 \end{cases}$, $\begin{cases} x + 3y + 2z = 4 \\ 6x + 2y - 4z = 3 \end{cases}$

En los ejercicios 83–85, verifique que las rectas dadas son perpendiculares (es decir, sus vectores paralelos son perpendiculares).

83. $\begin{cases} x = 2 - 3t \\ y = 3 - t \\ z = 1 + t \end{cases}$, $\begin{cases} x = 5 + t \\ y = 1 + t \\ z = 5 + 4t \end{cases}$

84. $\begin{cases} x = 2 - t \\ y = 3 - t \\ z = 5 + t \end{cases}$, $\begin{cases} x + y - z = 5 \\ 2x - 4z = 7 \end{cases}$

85. $\begin{cases} 2x - y - z = 7 \\ x + y + z = 6 \end{cases}$, $\begin{cases} x + y + z = 4 \\ x + 3y - 5z = 5 \end{cases}$

(*) **86.** Demuestre que la distancia d entre el punto $\mathbf{p} = (x_1, y_1, z_1)$ y la recta $x = x_0 + at$, $y = y_0 + bt$, $z = z_0 + ct$, está dada por

$$d = \frac{\|(a, b, c) \times (x_1 - x_0, y_1 - y_0, z_1 - z_0)\|}{\|(a, b, c)\|}$$

En los ejercicios 87–89, use el resultado del ejercicio anterior para calcular la distancia entre el punto $\mathbf{p}$ y la recta dada.

87. $\mathbf{p} = (1, 2, 3)$, $\begin{cases} x = 5 - 4t \\ y = 1 + t \\ z = 3t \end{cases}$

88. $\mathbf{p} = (-2, 4, 5)$, $\begin{cases} x = 1 + t \\ y = 2 + 3t \\ z = -1 + 5t \end{cases}$

89. $\mathbf{p} = (1, 1, 1)$, $\begin{cases} x + y - 2z = 5 \\ x - y - z = 13 \end{cases}$

En los ejercicios 90–92, calcule la distancia entre las dos rectas dadas (ver ejemplo 9).

90. $\begin{cases} x = 1 + 2t \\ y = -1 + t \\ z = t \end{cases}$, $x = y = 2z$

91. $\begin{cases} x = 5 + 3t \\ y = 1 - 2t, \\ z = 3 \end{cases}$ $\begin{cases} x = 1 \\ y = 5t \\ z = 2 - t \end{cases}$

92. $x = y = z$ $\begin{cases} 3x - y + z = 5 \\ x + y + z = 4 \end{cases}$

(*) 93. **(Planos en $\mathbb{R}^n$).** Dado un punto $\mathbf{p}_0 = (x_{1_0}, x_{2_0}, \ldots, x_{n_0}) \in \mathbb{R}^n$ y un vector no nulo $\mathbf{v} = (a_1, a_2, \ldots, a_n)$, el *plano* (o hiperplano) en $\mathbb{R}^n$ que pasa por $\mathbf{p}_0$ y tiene a $\mathbf{v}$ por vector normal es el conjunto de puntos $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ tales que $\mathbf{v} \cdot (\mathbf{x} - \mathbf{p}_0) = 0$, o bien, tales que

$$a_1(x_1 - x_{1_0}) + a_2(x_2 - x_{2_0}) + \cdots + a_n(x_n - x_{n_0}) = 0$$

a. Concilie esta definición con la dada en el texto.

b. Encuentre la ecuación del plano en $\mathbb{R}^4$ que pase por el origen y tenga al vector $\mathbf{v} = (1, 1, -1, -1)$ por vector normal.

c. Encuentre la ecuación del plano en que pase por el punto $\mathbf{p}_0 = (1, 2, 0, -1)$ y tenga al vector $\mathbf{v} = (0, 1, 0, 1)$ por vector normal.

d. Dos planos en $\mathbb{R}^n$ se dicen ser *paralelos* si sus vectores normales son linealmente dependientes. Encuentre la ecuación del plano en $\mathbb{R}^4$ paralelo al plano $x + 3y - 6z + 5u = 3$ que pase por el punto $\mathbf{p}_0 = (1, 1, 1, 0)$.

e. Dos planos en $\mathbb{R}^n$ se dicen ser *ortogonales* si sus vectores normales lo son. Si denotamos por (x, y, z, u) los puntos de $\mathbb{R}^4$, demuestre que los planos coordenados $x = 0$, $y = 0$, $z = 0$, $u = 0$ son ortogonales entre sí. Más aún, (demuestre que) los planos coordenados del espacio $\mathbb{R}^n$, a saber $x_1 = 0, x_2 = 0, \ldots, x_n = 0$, son planos ortogonales entre sí.

f. Considere los siguientes cuatro planos en $\mathbb{R}^4$

Π_1: pasa por $\mathbf{p}_0 = (1, 2, 0, 2)$ y tiene a $\mathbf{v} = (1, 1, 1, 1)$ por vector normal.
Π_2: pasa por $\mathbf{p}_0 = (0, 0, 0, 0)$ y tiene a $\mathbf{v} = (2, -1, 1, 3)$ por vector normal.
Π_3: pasa por $\mathbf{p}_0 = (3, 0, -2, -1)$ y tiene a $\mathbf{v} = (-1, 4, -2, -3)$ por vector normal.
Π_4: pasa por $\mathbf{p}_0 = (1, 1, -1, 0)$ y tiene a $\mathbf{v} = (1, -3, 1, 1)$ por vector normal.

Demuestre que estos planos tienen un punto común. Determínelo.

g. Determine la ecuación del plano en $\mathbb{R}^4$ que pasa por el origen de coordenadas y por los puntos $A = (1, 0, 2, 1)$, $B = (1, 1, 0, 0)$, $C = (0, 2, 1, 3)$.

h. Determine la ecuación del plano en $\mathbb{R}^5$ que pasa por el origen de coordenadas y por los puntos $A = (1, 0, 1, 0, 0)$, $B = (3, 1, 1, 2, 1)$, $C = (0, 1, -1, 2, 3)$, $D = (1, 1, -1, -1, -1)$.

i. Determine la ecuación del plano en $\mathbb{R}^n$ que pasa por los n puntos $\mathbf{p}_i = (0, \ldots, 0, a_i, 0, \ldots, 0)$ (el número $a_i \neq 0$ en la i-ésima coordenada), $i = 1, 2, \ldots, n$.

(*) 94. **(Rectas en $\mathbb{R}^n$).** Dado un punto $\mathbf{p}_0 = (x_{1_0}, x_{2_0}, \ldots, x_{n_0}) \in \mathbb{R}^n$ y un vector no nulo $\mathbf{v} = (a_1, a_2, \ldots, a_n)$, la *recta* en $\mathbb{R}^n$ que pasa por $\mathbf{p}_0$ y tiene a $\mathbf{v}$ por vector paralelo es el conjunto de puntos $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ tales que el vector $\mathbf{x} - \mathbf{p}_0$ es paralelo al vector $\mathbf{v}$, o bien, en forma más explícita, tales que $x_i = x_{i_0} + a_i t$, donde $t \in \mathbb{R}$, $i = 1, 2, \ldots, n$. Obsérvese que la ecuación de una recta semejante se puede escribir como una serie de $n - 1$ igualdades de la siguiente manera

$$\frac{x_1 - x_{1_0}}{a_1} = \frac{x_2 - x_{2_0}}{a_2} = \cdots = \frac{x_n - x_{n_0}}{a_n}$$

a. Determine la ecuación de la recta en $\mathbb{R}^4$ que pasa por el origen de coordenadas y tiene al vector $\mathbf{v} = (1, 1, 1, 1)$ por vector paralelo.

b. Encuentre la ecuación de la recta en $\mathbb{R}^4$ que pasa por el punto $\mathbf{p}_0 = (2, 1, 1, 4)$ y tiene al vector $\mathbf{v} = (-1, 1, 3, -5)$ por vector paralelo.

c. Encuentre la ecuación de la recta en $\mathbb{R}^4$ que pasa por el punto $\mathbf{p}_0 = (1, -1, 3, -4)$ y es ortogonal al plano $3x - 2y + 4z - 3u = 7$ (es decir, los vectores paralelo a la recta y normal al plano, son paralelos).

d. Encuentre la ecuación de la recta en $\mathbb{R}^n$ que pasa por el origen de coordenadas y tiene al vector $\mathbf{v} = (1, 1, \ldots, 1)$ por vector paralelo.

e. Determine la ecuación de la recta en $\mathbb{R}^4$ que pasa por los dos puntos $\mathbf{p} = (2, 1, 1, 3)$, $\mathbf{q} = (3, 1, 1, 6)$.

f. Dos rectas en $\mathbb{R}^n$ se dicen ser paralelas si sus vectores paralelos correspondientes lo son. Determine la ecuación de la recta en $\mathbb{R}^5$ que pasa por el punto $\mathbf{p} = (1, 0, 0, 0, 1)$ y es paralela a la recta

$$\frac{x_1 - 2}{3} = \frac{x_2 - 5}{4} = \frac{x_3 + 4}{-3} = \frac{x_4 - 9}{1} = \frac{x_5 - 7}{7}$$

g. Determine el punto donde la recta en $\mathbb{R}^4$

$$\frac{x_1 - 3}{10} = \frac{x_2 - 1}{10} = \frac{x_3 - 2}{-1} = \frac{x_4 - 5}{-1}$$

intersecta al plano $x_1 + 3x_2 + 5x_3 - 2x_4 = 80$.

h. Demuestre que los siguientes tres planos en $\mathbb{R}^4$: $x + y - 2z - u = 2$, $3x - y + z + u = 1$, $x - y - z + 2u = 1$, se intersectan en los puntos de la recta en $\mathbb{R}^4$ que pasa por $(0, -10, -3, -6)$ y tiene al vector $(1, 17, 4, 10)$ por vector paralelo.

i. Demuestre que las dos rectas en $\mathbb{R}^4$

$$L_1\colon \begin{cases} x = 2t \\ y = t - 3 \\ z = 3t - 3 \\ u = 2t - 3 \end{cases}, \qquad \begin{cases} x = 2t \\ y = -3t + 2 \\ z = t + 2 \\ u = t \end{cases}, \qquad t \in \mathbb{R}$$

se intersectan en un punto. Determínelo.

j. Considere el triángulo en $\mathbb{R}^4$ cuyos vértices son $A = (1, 1, 2, 1)$, $B = (3, 1, 3, -2)$, $C = (4, -3, 2, 2)$. Determine las ecuaciones de las medianas de este triángulo. Compruebe que éstas se cortan en un punto.

1.7 Transformaciones lineales

En esta sección vamos a recordar algunas de las definiciones y resultados básicos que aparecen en el estudio de las transformaciones lineales, (tema que pertenece al ámbito del Álgebra Lineal). Comenzamos por dar la definición de este tipo de funciones.

Definición. Se dice que la función $T\colon \mathbb{R}^n \to \mathbb{R}^m$ es una *transformación lineal* si: (1) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$; (2) $T(c\mathbf{x}) = cT(\mathbf{x})$, $c \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$. El *núcleo* de la transformación T, denotado por Ker T, es el conjunto de los vectores $\mathbf{x} \in \mathbb{R}^n$ cuya imagen

es el vector $\mathbf{0}$ de $\mathbb{R}^m$, y la *imagen* de la transformación T, denotada por Im T, es el conjunto de los vectores $\mathbf{y} \in \mathbb{R}^m$ tales que $\mathbf{y} = T(\mathbf{x})$ para alguna $\mathbf{x} \in \mathbb{R}^n$. Es decir

$$\operatorname{Ker} T = \{\mathbf{x} \in \mathbb{R}^n | T(\mathbf{x}) = \mathbf{0}\}$$
$$\operatorname{Im} T = \{\mathbf{y} \in \mathbb{R}^m | \mathbf{y} = T(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n\}$$ ■

A una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^n$ del espacio $\mathbb{R}^n$ en él mismo se le llama *operador lineal* (en $\mathbb{R}^n$). Una condición equivalente para que la función $T: \mathbb{R}^n \to \mathbb{R}^m$ sea una transformación lineal, que contiene a las condiciones (1) y (2) de la definición anterior, es que $T(\mathbf{x} + c\mathbf{y}) = T(\mathbf{x}) + cT(\mathbf{y})$, $c \in \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Obsérvese que siendo $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, entonces $T(\mathbf{0}) = \mathbf{0}$. Es decir, la imagen del vector $\mathbf{0} \in \mathbb{R}^n$ es siempre el vector $\mathbf{0} \in \mathbb{R}^m$. En efecto, usando la propiedad (1) de la definición anterior, se tiene $T(\mathbf{x}) = T(\mathbf{x} + \mathbf{0}) = T(\mathbf{x}) + T(\mathbf{0})$, de donde, por la unicidad del vector $\mathbf{0} \in \mathbb{R}^m$ se tiene que $T(\mathbf{0}) = \mathbf{0}$. Dejamos que el lector dé otro argumento de este hecho utilizando la propiedad (2) de la definición dada de transformación lineal. Se tiene entonces que tanto el núcleo como la imagen de una transformación lineal son conjuntos no vacíos (de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente), pues al menos el vector $\mathbf{0}$ de $\mathbb{R}^n$ está en su núcleo y al menos el vector $\mathbf{0}$ de $\mathbb{R}^m$ está en su imagen.

Otro hecho importante que se deduce fácilmente de la definición dada de transformación lineal es que tanto el núcleo de T como su imagen, son *subespacios* (de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente). En efecto, si $\mathbf{x}, \mathbf{x}' \in \operatorname{Ker} T$ (i.e. $T(\mathbf{x}) = T(\mathbf{x}') = \mathbf{0}$), entonces $T(\mathbf{x} + \mathbf{x}') = T(\mathbf{x}) + T(\mathbf{x}') = \mathbf{0} + \mathbf{0} = \mathbf{0}$, por lo que $\mathbf{x} + \mathbf{x}'$ pertenece también a Ker T. Por otra parte, si $c \in \mathbb{R}$ y $\mathbf{x} \in \operatorname{Ker} T$, entonces $T(c\mathbf{x}) = cT(\mathbf{x}) = c(\mathbf{0}) = \mathbf{0}$, por lo que $c\mathbf{x}$ también en Ker T. Según el teorema 1.1.1 concluimos entonces que Ker T es un subespacio de $\mathbb{R}^n$. Dejamos al lector verificar que Im T es también un subespacio de $\mathbb{R}^m$ (codominio de la transformación T). A la dimensión del núcleo de la transformación lineal T se le llama *nulidad de T*, y a la dimensión de su imagen se le llama *rango de T*.

Uno de los resultados clásicos importantes que aparecen en el estudio de las transformaciones lineales $T: \mathbb{R}^n \to \mathbb{R}^m$, que relaciona la nulidad y el rango de T con n = dimensión del espacio $\mathbb{R}^n$, es el siguiente teorema que enunciamos sin demostración.

Teorema 1.7.1 (Teorema de la dimensión). Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces

$$\dim(\operatorname{Ker} T) + \dim(\operatorname{Im} T) = n$$ ■

Ejemplo 1. La función identidad $T: \mathbb{R}^n \to \mathbb{R}^n$, $T(\mathbf{x}) = \mathbf{x}$, es una transformación lineal, pues $T(\mathbf{x} + c\mathbf{y}) = \mathbf{x} + c\mathbf{y} = T(\mathbf{x}) + cT(\mathbf{y})$, $c \in \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Su núcleo consta solamente del vector $\mathbf{0}$ de $\mathbb{R}^n$ y su imagen es todo $\mathbb{R}^n$. La nulidad de T es entonces 0 y su rango es n. ■

Ejemplo 2. La función $T: \mathbb{R}^2 \to \mathbb{R}^3$, dada por $T(x, y) = (x+y, 2x, 3x-4y)$ es una transformación lineal, pues

$$\begin{aligned}
T((x, y) + c(x', y')) &= T(x + cx', y + cy') \\
&= (x + cx' + y + cy', 2(x + cx'), 3(x + cx') - 4(y + cy')) \\
&= (x + y + c(x' + y'), 2x + c(2x'), 3x - 4y + c(3x' - 4y')) \\
&= (x + y, 2x, 3x - 4y) + c(x' + y', 2x', 3x' - 4y') \\
&= T(x, y) + cT(x', y')
\end{aligned}$$

Su núcleo está formado por aquellos vectores (x, y) de $\mathbb{R}^2$ tales que

$$T(x, y) = (x + y, 2x, 3x - 4y) = (0, 0, 0)$$

Es claro que sólo el vector $(0, 0)$ de $\mathbb{R}^2$ cumple con esta condición. Es entonces el único vector del núcleo de la transformación T. La imagen de T está formada por aquellos vectores (x, y, z) de $\mathbb{R}^3$ para los cuales existe $(a, b) \in \mathbb{R}^2$ tal que $T(a, b) = (x, y, z)$. Es decir, vectores (x, y, z) para los que existe a y b tales que

$$T(a, b) = (a + b, 2a, 3a - 4b) = (x, y, z)$$

Esta condición es equivalente a plantear el siguiente sistema de 3 ecuaciones en las indeterminadas a, b

$$a + b = x, \qquad 2a = y, \qquad 3a - 4b = z$$

del cual queremos saber condiciones sobre x, y, z para que el sistema tenga solución. Usando eliminación en este sistema, se llega a que (x, y, z) está en la imagen de T si y sólo si $8x - 7y + 2z = 0$. Es decir, la imagen de T está formada por los vectores del plano $8x - 7y + 2z = 0$. El lector puede verificar que $\beta = \{(1, 0, -4), (0, 2, 7)\}$ es una base de Im T, por lo que el rango de T es 2. Se cumple entonces el teorema de la dimensión, pues la nulidad de T (que es 0) sumada al rango de T es igual a $2 = \dim \mathbb{R}^2$. ■

Una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ queda completamente determinada si conocemos las imágenes de los vectores $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ de una base de $\mathbb{R}^n$. Más aún, si $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base del espacio $\mathbb{R}^n$, existe una única transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ tal que $T(\mathbf{e}_i) = \mathbf{y}_i \in \mathbb{R}^m$, en la que $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ son vectores dados. En efecto, si $\mathbf{x} \in \mathbb{R}^n$ es un vector cualquiera de $\mathbb{R}^n$, podemos escribirlo en términos de la base β como

$$\mathbf{x} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \ldots + c_n\mathbf{e}_n$$

donde $c_1, c_2, \ldots, c_n$ son escalares bien determinados. Entonces la imagen de $\mathbf{x}$ bajo T queda determinada por (aplicando el hecho de que T es *lineal*)

$$\begin{aligned} T(\mathbf{x}) &= T(c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n) \\ &= c_1T(\mathbf{e}_1) + c_2T(\mathbf{e}_2) + \cdots + c_nT(\mathbf{e}_n) \\ &= c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_n\mathbf{y}_n \end{aligned}$$

Dejamos a cargo del lector la verificación de que esta función es efectivamente una transformación lineal y que es única.

Ejemplo 3. Según la discusión anterior, sólo hay una transformación lineal $T: \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(1, 0) = (1, 2, 3)$, $T(0, 1) = (1, 0, -4)$. Esta transformación es tal que en (x, y) es

$$\begin{aligned} T(x, y) &= T(x(1, 0) + y(0, 1)) = xT(1, 0) + yT(0, 1) = x(1, 2, 3) + y(1, 0, -4) \\ &= (x + y, 2x, 3x - 4y) \end{aligned}$$

Esta es, de hecho, la transformación lineal del ejemplo 2. ■

Una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ es inyectiva (es decir, si $\mathbf{x}_1$ y $\mathbf{x}_2 \in \mathbb{R}^n$ son dos vectores *distintos*, entonces sus imágenes $T(\mathbf{x}_1)$ y $T(\mathbf{x}_2)$ son *distintas*) si y sólo si su núcleo consta solamente

del vector $\mathbf{0} \in \mathbb{R}^n$. En efecto, suponga que Ker $T = \{\mathbf{0}\}$ y que $T(\mathbf{x}_1) = T(\mathbf{x}_2)$; siendo T lineal se tiene $T(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, de donde $\mathbf{x}_1 - \mathbf{x}_2$ es un vector del núcleo de T; debe ser entonces el vector $\mathbf{0}$; se debe tener entonces que $\mathbf{x}_1 = \mathbf{x}_2$. Esto prueba que T es inyectiva. Por otra parte, si T es inyectiva, y el vector $\mathbf{v}$ está en el núcleo de T, entonces $T(\mathbf{0}) = T(\mathbf{v}) = \mathbf{0}$, por lo que $\mathbf{v} = \mathbf{0}$ (¡pues T es inyectiva!); es decir, Ker $T = \{\mathbf{0}\}$. Esto prueba nuestra afirmación.

Si el operador lineal $T: \mathbb{R}^n \to \mathbb{R}^n$ es inyectivo, se dice que es un *isomorfismo.* De hecho, se puede probar que T es un isomorfismo si y sólo si T es sobreyectivo (es decir, si su imagen es todo $\mathbb{R}^n$).

Un hecho que usaremos de forma importante en el capítulo 2 de este libro es que a toda transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ se le puede asociar una matriz de orden $m \times n$, que denotaremos por $[T]$, tal que la acción de T sobre un vector $\mathbf{x} \in \mathbb{R}^n$ pueda verse como "una multiplicación por la matriz $[T]$". Los detalles son como sigue: fijamos inicialmente una base $\beta_1 = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$ y una base $\beta_2 = \{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_m\}$ de $\mathbb{R}^m$. Para simplificar la discusión, pensemos en estas bases como las canónicas. Vamos siempre a identificar a un vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p) \in \mathbb{R}^p$ con su matriz de coordenadas (respecto de la base canónica)

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_p \end{bmatrix}$$

Dado un vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, la imagen de él bajo T es un vector de $\mathbb{R}^m$, digamos $T(\mathbf{x}) = (c_1, c_2, \ldots, c_m)$. Por otra parte, con el hecho de que T es lineal y escribiendo al vector $\mathbf{x}$ en términos de la base β_1, se tiene

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= \sum_{j=1}^{n} x_jT(\mathbf{e}_j) \end{aligned}$$

Cada vector $T(\mathbf{e}_j) \in \mathbb{R}^m$ se puede escribir en términos de la base $\beta_2 = \{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_m\}$ como

$$T(\mathbf{e}_j) = a_{1j}\mathbf{e}'_1 + a_{2j}\mathbf{e}'_2 + \cdots + a_{mj}\mathbf{e}'_m = \sum_{i=1}^{m} a_{ij}\mathbf{e}'_i \qquad j = 1, 2, \ldots, n$$

Entonces

$$T(\mathbf{x}) = \sum_{j=1}^{n} x_jT(\mathbf{e}_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\mathbf{e}'_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) \mathbf{e}'_i$$

En vista de la unicidad de la representación de un vector en términos de una base, concluimos que, siendo $T(\mathbf{x}) = (c_1, c_2, \ldots, c_m)$, los escalares c_i $i = 1, 2, \ldots, m$ son

$$c_i = \sum_{j=1}^{n} a_{ij}x_j$$

En otras palabras, el vector $T(\mathbf{x})$ se puede presentar como

$$T(\mathbf{x}) = A\mathbf{x}$$

donde la matriz A (que denotaremos por $[T]$), de orden $m \times n$ tiene elementos a_{ij}. Decimos que esta matriz *representa* a la transformación T (en términos de las bases canónicas de los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$). Más aún, la j-ésima columna de esta matriz tiene las coordenadas del vector $T(\mathbf{e}_j)$ en términos de la base β_2. Esquemáticamente

$$[T] = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

Ejemplo 4. Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (2x - y, 5x + 8y)$. Tomando la base canónica de $\mathbb{R}^2$ tenemos que $T(\mathbf{e}_1) = T(1, 0) = (2, 5)$, $T(\mathbf{e}_2) = T(0, 1) = (-1, 8)$. Estos vectores van a constituir las columnas de la matriz $[T]$. De hecho

$$[T] = \begin{bmatrix} 2 & -1 \\ 5 & 8 \end{bmatrix}$$

Obsérvese entonces que la imagen del vector (x, y) se puede determinar como

$$T(x, y) = \begin{bmatrix} 2 & -1 \\ 5 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - y \\ 5x + 8y \end{bmatrix}$$ ■

Ejemplo 5. Consideremos la transformación lineal $T: \mathbb{R}^2 \to \mathbb{R}^3$ del ejemplo 2, $T(x, y) = (x + y, 2x, 3x - 4y)$. Puesto que

$$T(1, 0) = (1, 2, 3), T(0, 1) = (1, 0, -4)$$

la matriz que representa a T es la matriz 3×2 dada por

$$[T] = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 3 & -4 \end{bmatrix}$$ ■

Dadas dos transformaciones lineales $T_1: \mathbb{R}^n \to \mathbb{R}^m$ y $T_2: \mathbb{R}^m \to \mathbb{R}^p$, podemos formar su composición $T_2 \circ T_1: \mathbb{R}^n \to \mathbb{R}^p$. Es inmediato verificar que ésta es también una transformación lineal. Más aún, haciendo uso de las matrices que representan a T_1 y T_2 podemos escribir

$$(T_2 \circ T_1)(\mathbf{x}) = T_2(T_1(\mathbf{x})) = T_2([T_1]\mathbf{x}) = [T_2][T_1]\mathbf{x}$$

de donde se ve que la matriz que representa a la composición $T_2 \circ T_1$ (la cual es una matriz de orden $p \times n$) es el producto de la matriz que representa a T_2 (que es de orden $p \times m$) por la matriz que representa a T_1 (que es de orden $m \times n$). O sea

$$[T_2 \circ T_1] = [T_2][T_1]$$

Consideremos el operador lineal $T: \mathbb{R}^n \to \mathbb{R}^n$. Es importante estudiar condiciones bajo las cuales este operador tiene un *inverso* $T^{-1}: \mathbb{R}^n \to \mathbb{R}^n$, en el sentido de que las composiciones $T \circ T^{-1}$ y $T^{-1} \circ T$ sean la función identidad de $\mathbb{R}^n$. Como cualquier función, es claro que para que exista T^{-1}, T debe ser inyectivo. En tal caso se dice que el operador T es *inversible* y es fácil comprobar

que T^{-1} es también un *operador lineal* (y así, su composición con T será entonces el *operador lineal* identidad de $\mathbb{R}^n$). Más aún, siendo T inversible (y solamente en este caso), la matriz $[T]$ que lo representa también es una matriz inversible , la matriz que representa a T^{-1} será (como todo el mundo esperaría que ocurriera) la inversa de la matriz que representa a T. Es decir, se tiene la bonita fórmula

$$[T^{-1}] = [T]^{-1}$$

(la cual es una consecuencia inmediata de que $I = [\text{Id}] = [T_2 \circ T_1] = [T_2][T_1]$). En el siguiente teorema, que enunciamos sin demostración, se establecen varias condiciones equivalentes al hecho de que el operador lineal $T: \mathbb{R}^n \to \mathbb{R}^n$ sea inversible.

Teorema 1.7.2 Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ un operador lineal en $\mathbb{R}^n$. Las siguientes afirmaciones son equivalentes:

1. T es inversible.
2. T es inyectivo.
3. $\text{Ker}\, T = \{\mathbf{0}\}$
4. $\text{Im}\, T = \mathbb{R}^n$
5. T es sobreyectivo.
6. La matriz $[T]$ es inversible.
7. $\det[T] \neq 0$.
8. El sistema de ecuaciones lineales $[T]\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.

Ejercicios (Capítulo 1, Sección 7)

En los ejercicios 1–5, constate que la función dada es una transformación lineal. En cada caso describa su núcleo y su imagen, y compruebe que se satisface el teorema de la dimensión.

1. $T: \mathbb{R}^2 \to \mathbb{R}$, $T(x, y) = x + y$
2. $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x, y) = (x + y, 2x - 3y)$
3. $T: \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (x + y, 3x - z)$
4. $T: \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (x - y, y - z, z - x)$
5. $T: \mathbb{R}^4 \to \mathbb{R}^2$, $T(x, y, z, u) = (x + y, z + u)$
6. Describa la estructura general de las transformaciones lineales cuyo dominio y codominio es $\mathbb{R}$.
7. Verifique que la función $T: \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (x - y, z + 1)$ no es una transformación lineal.
8. Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal tal que $T(1, 0) = (2, 5)$, $T(0, 1) = (-3, 4)$. Halle una expresión general para $T(x, y)$.
9. Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal tal que $T(1, 1) = (3, 2)$, $T(0, 1) = (2, 7)$. Halle una expresión general para $T(x, y)$.
10. Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ un operador lineal cuyo núcleo es $\text{Ker}\, T = \{(x, y) | x = y\}$. Demuestre que T no es sobreyectiva. Más aún, pruebe que $\text{Im}\, T$ es una recta que pasa por el origen. (Sugerencia: use el teorema de la dimensión).

11. Sea $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ un operador lineal cuya imagen es $\operatorname{Im} T = \{(x, y, z) | x = y = z\}$. Demuestre que el núcleo de T es un plano en $\mathbb{R}^3$ que pasa por el origen.

12. Sea $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ una transformación lineal tal que $T(1, 1, 0) = T(1, 1, 1) = (0, 0)$, $T(2, 3, -1) = (a, b) \neq (0, 0)$. Demuestre que el núcleo de T es un plano en $\mathbb{R}^3$ que pasa por el origen. Halle su ecuación.

13. Demuestre que el operador lineal $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es inyectivo si y sólo si es sobreyectivo.

14. Sean $T_1: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ y $T_2: \mathbb{R}^m \longrightarrow \mathbb{R}^p$ dos transformaciones lineales. Demuestre que su composición $T_2 \circ T_1: \mathbb{R}^n \longrightarrow \mathbb{R}^p$ es una transformación lineal.

15. Sean $T_1: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $T_2: \mathbb{R}^m \longrightarrow \mathbb{R}^p$, $T_3: \mathbb{R}^p \longrightarrow \mathbb{R}^q$ transformaciones lineales. Demuestre la propiedad asociativa de su composición. Es decir, demuestre que las transformaciones lineales $T_3 \circ (T_2 \circ T_1)$, $(T_3 \circ T_2) \circ T_1: \mathbb{R}^n \longrightarrow \mathbb{R}^q$ son iguales. ¿Es esta propiedad válida para funciones en general?

16. Sean $T_1: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, T_2, $T_3: \mathbb{R}^m \longrightarrow \mathbb{R}^p$ transformaciones lineales. Demuestre la propiedad distributiva de la composición respecto de la suma: $T_1 \circ (T_2 + T_3) = T_1 \circ T_2 + T_1 \circ T_3$. ¿Es esta propiedad válida para funciones en general?

17. Sea $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ un operador lineal inyectivo. Demuestre que el inverso $T^{-1}: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ (tal que $T^{-1} \circ T$ y $T \circ T^{-1}$ es la identidad en $\mathbb{R}^n$) es un operador lineal. (Esbozo de la demostración: Sea $T(\mathbf{x}) = \mathbf{u}$, $T(\mathbf{y}) = \mathbf{v}$. Se tiene que $T^{-1}(\mathbf{u} + c\mathbf{v}) = T^{-1}(T(\mathbf{x}) + cT(\mathbf{y})) = T^{-1}(T(\mathbf{x} + c\mathbf{y})) = (T^{-1} \circ T)(\mathbf{x} + c\mathbf{y}) = \mathbf{x} + c\mathbf{y} = T^{-1}(\mathbf{u}) + cT^{-1}(\mathbf{v})$).

Para cada uno de los operadores lineales de los ejercicios 18–22, halle la matriz que lo representa (en las bases canónicas de los espacios correspondientes). En base a ella, argumente por qué el operador dado es inversible. Determine una fórmula explícita para el operador inverso T^{-1}.

18. $T: \mathbb{R} \longrightarrow \mathbb{R}$, $T(x) = ax$, $a \neq 0$.

19. $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $T(x, y) = (x + y, x - y)$.

20. $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $T(x, y) = (10x + 6y, x - 3y)$.

21. $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $T(x, y, z) = (x, x + y, x + y + z)$

22. $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $T(x, y, z) = (z, x, y)$

23. Considere la transformación lineal $T_2 \circ T_1: \mathbb{R}^n \longrightarrow \mathbb{R}^p$ del ejercicio 14. Demuestre que $\operatorname{Ker} T_1 \subseteq \operatorname{Ker} T_2 \circ T_1$, $\operatorname{Im} T_2 \circ T_1 \subseteq \operatorname{Im} T_2$.

24. Sean $T_1, T_2: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ dos operadores lineales. Demuestre que si T_2 es inyectivo, entonces $\operatorname{Ker} T_1 = \operatorname{Ker} T_2 \circ T_1$, y si T_1 es inyectivo, entonces $\operatorname{Im} T_2 = \operatorname{Im} T_2 \circ T_1$.

25. Considere las transformaciones lineales $T_1: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, $T_1(x, y, z) = (x + y, x - z)$, $T_2: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$, $T_2(x, y) = (y, x, x + y)$. Determine expresiones explícitas para las transformaciones lineales $T_1 \circ T_2: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ y $T_2 \circ T_1: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$. Verifique en cada caso que la matriz que representa a la composición de transformaciones es el producto de las matrices que representan a cada una de las transformaciones lineales que se componen.

26. Repita el ejercicio anterior con los operadores lineales $T_1, T_2: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $T_1(x, y, z) = (x - z, y, x)$, $T_2(x, y, z) = (y + z, x + z, z)$.

27. Verifique que el operador lineal $T_2 \circ T_1: \mathbb{R}^3 \to \mathbb{R}^3$ obtenido en el ejercicio 25 no es inversible.

(*) **28.** Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Demuestre que si $n > m$, T no puede ser inyectiva. ¿Puede ser sobreyectiva?

(*) **29.** Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Demuestre que si $n < m$, T no puede ser sobreyectiva. ¿Puede ser inyectiva?

30. Demuestre que una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ con $n \neq m$, no puede ser inversible.

(*) **31.** Sean $T_1: \mathbb{R}^n \to \mathbb{R}^m$, $T_2: \mathbb{R}^m \to \mathbb{R}^n$ transformaciones lineales. Demuestre que si $n > m$, el operador lineal $T_2 \circ T_1: \mathbb{R}^n \to \mathbb{R}^n$ no puede ser inversible. ¿Puede ser inversible el operador $T_1 \circ T_2: \mathbb{R}^m \to \mathbb{R}^m$? Explique.

(*) **32.** **(Aplicaciones bilineales)** Una *aplicación bilineal* (en $\mathbb{R}^n$) es una función $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ tal que: 1. $f(\mathbf{u}_1 + c\mathbf{u}_2, \mathbf{v}) = f(\mathbf{u}_1, \mathbf{v}) + cf(\mathbf{u}_2, \mathbf{v})$, para cada $\mathbf{v} \in \mathbb{R}^n$ (fija) y todo $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^n$, $c \in \mathbb{R}$. 2. $f(\mathbf{u}, \mathbf{v}_1 + c\mathbf{v}_2) = f(\mathbf{u}, \mathbf{v}_1) + cf(\mathbf{u}, \mathbf{v}_2)$, para cada $\mathbf{u} \in \mathbb{R}^n$ (fija) y todo $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, $c \in \mathbb{R}$. Obsérvese entonces que una aplicación bilineal es aquélla que, fijando una de sus variables, se comporta de manera lineal respecto de la otra. Cuando $m = 1$, se dice que f es una *forma* bilineal.

a. Verifique que la función $f: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f(x, y) = x + y$ es una forma bilineal en $\mathbb{R}$. ¿Es una transformación lineal de $\mathbb{R}^2 (\approx \mathbb{R} \times \mathbb{R})$ a $\mathbb{R}$?

b. Verifique que la función $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $f\big((x_1, y_1), (x_2, y_2)\big) = x_1x_2 + y_1y_2$ es una forma bilineal en $\mathbb{R}^2$.

c. Demuestre que el producto punto $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ es una forma bilineal en $\mathbb{R}^n$.

d. Demuestre que el producto cruz $\times: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ es una aplicación bilineal en $\mathbb{R}^3$.

e. Demuestre que si $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación bilineal, entonces $f(\mathbf{0}, \mathbf{v}) = f(\mathbf{u}, \mathbf{0}) = \mathbf{0}$, $\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

(*) **33.** **(Aplicaciones m-lineales).** Una *aplicación m-lineal* (en $\mathbb{R}^n$) es una función $f: \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^p$ (definida en el producto cartesiano de $\mathbb{R}^n$ consigo mismo m veces) tal que

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i + c\mathbf{v}_i', \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) = f(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n) + cf(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i', \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n)$$

para todo $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ (fijos) y todo $\mathbf{v}_i, \mathbf{v}_i' \in \mathbb{R}^n$, $c \in \mathbb{R}$, $i = 1, 2, \ldots, n$.

a. Demuestre que la función $f: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{u}, \mathbf{v}, \mathbf{w}) = a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$, en donde a, b, c son números reales dados, es una aplicación trilineal en $\mathbb{R}^n$.

b. Demuestre que si $f: \mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}^q$ es una aplicación m-lineal en $\mathbb{R}^n$, entonces $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{0}, \ldots, \mathbf{v}_n) = \mathbf{0}$.

(*) **34.** **(Aplicaciones m-lineales alternadas)** Se dice que la aplicación m-lineal en $\mathbb{R}^n$, $f: \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^p$ es *alternada,* si

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_n) = -f(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n)$$

(es decir, si al intercambiar de posición dos de sus argumentos, la imagen "cambia de signo").

a. Demuestre que si $f: \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^p$ es una aplicación m-lineal alternada en $\mathbb{R}^n$, entonces $f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n) = \mathbf{0}$.

b. Demuestre que si $f: \mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}^p$ es una aplicación m-lineal alternada en $\mathbb{R}^n$, entonces

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j + c\mathbf{v}_i, \ldots, \mathbf{v}_n) = f(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_n)$$

c. Demuestre que el producto cruz $\times: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ es una aplicación bilineal alternada en $\mathbb{R}^3$.

d. Sea A una matriz 2×2. Piense las columnas de esta matriz como las coordenadas de (dos) vectores en $\mathbb{R}^2$, digamos que si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, escribimos $\mathbf{v}_1 = (a, b)$, $\mathbf{v}_2 = (b, d)$. Demuestre que la función determinante $\det: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\det A = \det(\mathbf{v}_1, \mathbf{v}_2)$ es una forma bilineal alternada en $\mathbb{R}^2$.

e. Más en general, si A es una matriz $n \times n$, podemos verla constituida por (sus columnas) n vectores en $\mathbb{R}^n$, digamos $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Demuestre que la función determinante $\det: \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$, vista como función de las columnas de la matriz A, es una forma n-lineal alternada en $\mathbb{R}^n$. Demuestre que esto sigue siendo válido si cambiamos la palabra "columnas" por "líneas".

(**) **35.** **(El producto cruz en $\mathbb{R}^n$)** El producto cruz entre vectores fue estudiado en la sección 5. Esta operación se definió como una operación entre *dos* vectores del espacio $\mathbb{R}^3$. En el ejercicio anterior se demostró que ésta es una aplicación bilineal alternada en $\mathbb{R}^3$. En este ejercicio se presenta una generalización de esta operación, para $n - 1$ vectores en $\mathbb{R}^n$. Sean entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$, $n - 1$ vectores en $\mathbb{R}^n$. Se define el producto cruz de estos vectores, denotado por $\mathbf{v}_1 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_{n-1}$, como el vector $\mathbf{v} \in \mathbb{R}^n$ tal que para todo vector $\mathbf{u} \in \mathbb{R}^n$, se tiene $\mathbf{u} \cdot \mathbf{v} = \det(\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}) =$ determinante de la matriz $n \times n$ cuyas columnas son los n vectores $\mathbf{u}, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. Nótese entonces que el vector $\mathbf{v}$ tiene en su j-ésima coordenada a $\det(\mathbf{e}_j, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1})$. Más aún, si escribimos $\mathbf{v}_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$, $j = 1, 2, \ldots, n - 1$, y consideramos la matriz de orden $n \times (n - 1)$, $A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n-1}}$, entonces, usando la expansión de un determinante a lo largo de su primera línea, se tiene que $\det(\mathbf{e}_j, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}) = (-1)^{i+j} \det A(j)$, donde $A(j)$ es la matriz de orden $(n-1) \times (n-1)$ que proviene de A eliminando en esta última su j-ésima línea.

a. Demuestre que

$$\mathbf{v}_1 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_{n-1} = \sum_{j=1}^{n} (-1)^{1+j} \det A(j) \mathbf{e}_j$$

(donde $\mathbf{e}_j$ es el j-ésimo vector de la base canónica de $\mathbb{R}^n$)

b. Demuestre que la aplicación $f: \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}) = \mathbf{v}_1 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_{n-1}$ es una aplicación $(n - 1)$-lineal alternada. Concluya entonces que si los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1} \in \mathbb{R}^n$ son linealmente dependientes, entonces su producto cruz es cero.

c. Demuestre que el vector $\mathbf{v}_1 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_{n-1}$ es ortogonal a cada uno de los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$. Es decir, demuestre que

$$(\mathbf{v}_1 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_{n-1}) \cdot \mathbf{v}_j = 0 \qquad \forall j = 1, 2, \ldots, n - 1$$

d. Sean $\mathbf{v}_1 = (1, 0, 2, 1)$, $\mathbf{v}_2 = (1, 1, 0, 0)$, $\mathbf{v}_3 = (0, 2, 1, 3)$. Calcule $\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3$.

e. Obtenga la ecuación del hiperplano en $\mathbb{R}^4$ generado por los vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ del inciso anterior. (Ver inciso g del ejercicio 93 de la sección 6).

f. Sean $\mathbf{v}_1 = (1, 0, 1, 0, 0)$, $\mathbf{v}_2 = (3, 1, 1, 2, 1)$, $\mathbf{v}_3 = (0, 1, -1, 2, 3)$, $\mathbf{v}_4 = (1, 1, -1, -1, -1)$. Calcule $\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3 \times \mathbf{v}_4$.

g. Obtenga la ecuación del hiperplano en $\mathbb{R}^5$ generado por los vectores $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$ del inciso anterior. (Ver inciso **h** del ejercicio 93 de la sección 6).

h. En este inciso y en el próximo, se generalizan (al caso de vectores en $\mathbb{R}^4$) los resultados establecidos en el ejercicio 20 de la sección 5. Sean $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ vectores en $\mathbb{R}^4$. Demuestre la identidad de Lagrange

$$(\mathbf{u}_1 \times \mathbf{u}_2 \times \mathbf{u}_3) \cdot (\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3) = \det \begin{bmatrix} \mathbf{u}_1\mathbf{v}_1 & \mathbf{u}_1\mathbf{v}_2 & \mathbf{u}_1\mathbf{v}_3 \\ \mathbf{u}_2\mathbf{v}_1 & \mathbf{u}_2\mathbf{v}_2 & \mathbf{u}_2\mathbf{v}_3 \\ \mathbf{u}_3\mathbf{v}_1 & \mathbf{u}_3\mathbf{v}_2 & \mathbf{u}_3\mathbf{v}_3 \end{bmatrix}$$

i. Demuestre que

$$\mathbf{u}_3 \cdot \big(\mathbf{u}_1 \times \mathbf{u}_2 \times (\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3)\big) = -(\mathbf{u}_1 \times \mathbf{u}_2 \times \mathbf{u}_3) \cdot (\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3)$$

j. En los siguientes incisos se construye un argumento que finalizará con la obtención de una fórmula que generaliza, al caso de vectores en $\mathbb{R}^4$, a la fórmula establecida en el ejercicio 22 de la sección 5, en donde se vió que para $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^3$ se tiene que $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$. El argumento seguirá las mismas ideas que en el ejercicio mencionado. Demuestre que existen constantes α, β, $\gamma \in \mathbb{R}$ tales que

$$\mathbf{u}_1 \times \mathbf{u}_2 \times (\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 + \gamma\mathbf{v}_3$$

k. Tomando el producto punto en ambos miembros de la expresión anterior, con los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$, concluya que

$$(\mathbf{v}_1 \cdot \mathbf{u}_1)\alpha + (\mathbf{v}_2 \cdot \mathbf{u}_1)\beta + (\mathbf{v}_3 \cdot \mathbf{u}_1)\gamma = 0$$
$$(\mathbf{v}_1 \cdot \mathbf{u}_2)\alpha + (\mathbf{v}_2 \cdot \mathbf{u}_2)\beta + (\mathbf{v}_3 \cdot \mathbf{u}_2)\gamma = 0$$

l. Defina los vectores $\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^3$ como

$$\mathbf{x} = (\alpha, \beta, \gamma), \quad \mathbf{y}_1 = (\mathbf{v}_1 \cdot \mathbf{u}_1, \mathbf{v}_2 \cdot \mathbf{u}_1, \mathbf{v}_3 \cdot \mathbf{u}_1), \quad \mathbf{y}_2 = (\mathbf{v}_1 \cdot \mathbf{u}_2, \mathbf{v}_2 \cdot \mathbf{u}_2, \mathbf{v}_3 \cdot \mathbf{u}_2)$$

Verifique entonces que $\mathbf{x} \cdot \mathbf{y}_1 = 0$, $\mathbf{x} \cdot \mathbf{y}_2 = 0$, y que, por lo tanto, $\mathbf{x}$ es un vector en la dirección de $\mathbf{y}_1 \times \mathbf{y}_2$. Es decir, debe existir una constante k tal que $\mathbf{x} = k(\mathbf{y}_1 \times \mathbf{y}_2)$.

m. Verifique que el vector $\mathbf{y}_1 \times \mathbf{y}_2$ está dado por

$$\left(\det \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{v}_2 & \mathbf{u}_1 \cdot \mathbf{v}_3 \\ \mathbf{u}_2 \cdot \mathbf{v}_2 & \mathbf{u}_2 \cdot \mathbf{v}_3 \end{bmatrix}, -\det \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_1 \cdot \mathbf{v}_3 \\ \mathbf{u}_2 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_3 \end{bmatrix} \det \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_1 \cdot \mathbf{v}_2 \\ \mathbf{u}_2 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_2 \end{bmatrix}\right)$$

n. Use los resultados de los incisos **h** e **i** para demostrar que la constante k del inciso l) es igual a -1.

o. Concluya entonces que el vector $\mathbf{u}_1 \times \mathbf{u}_2 \times (\mathbf{v}_1 \times \mathbf{v}_2 \times \mathbf{v}_3)$ se puede escribir como

$$-\det\begin{bmatrix} \mathbf{u}_1\cdot\mathbf{v}_2 & \mathbf{u}_1\cdot\mathbf{v}_3 \\ \mathbf{u}_2\cdot\mathbf{v}_2 & \mathbf{u}_2\cdot\mathbf{v}_3 \end{bmatrix}\mathbf{v}_1 + \det\begin{bmatrix} \mathbf{u}_1\cdot\mathbf{v}_1 & \mathbf{u}_1\cdot\mathbf{v}_3 \\ \mathbf{u}_2\cdot\mathbf{v}_1 & \mathbf{u}_2\cdot\mathbf{v}_3 \end{bmatrix}\mathbf{v}_2 - \det\begin{bmatrix} \mathbf{u}_1\cdot\mathbf{v}_1 & \mathbf{u}_1\cdot\mathbf{v}_2 \\ \mathbf{u}_2\cdot\mathbf{v}_1 & \mathbf{u}_2\cdot\mathbf{v}_2 \end{bmatrix}\mathbf{v}_3$$

p. Para el caso de vectores en $\mathbb{R}^n$ se tiene el siguiente resultado que generaliza a la fórmula establecida en el ejercicio 22 de la sección 5: sean $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-2}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$ vectores en $\mathbb{R}^n$ (con $n \geq 3$). Entonces

$$\mathbf{u}_1 \times \mathbf{u}_2 \times \cdots \times \mathbf{u}_{n-2} \times (\mathbf{v}_1 \times \mathbf{v}_2 \times \cdots \times \mathbf{v}_{n-1}) = (-1)^{n+1}\sum_{i=1}^{n-1}\beta_i\mathbf{v}_i$$

en donde $\beta_i = \mathbf{e}_i \cdot (\alpha_1 \times \alpha_2 \times \ldots \times \alpha_{n-2})$, $i = 1, 2, \ldots, n-1$, siendo $\mathbf{e}_i$ el i-ésimo vector de la base canónica de $\mathbb{R}^{n-1}$ y

$$\alpha_j = (\mathbf{u}_j \cdot \mathbf{v}_1, \mathbf{u}_j \cdot \mathbf{v}_2, \ldots, \mathbf{u}_j \cdot \mathbf{v}_{n-1}) \qquad j = 1, 2, \ldots, n-2.$$

1.8 Valores y vectores propios

Sea $T\colon \mathbb{R}^n \to \mathbb{R}^n$ un operador lineal en $\mathbb{R}^n$. Se dice que el número real λ es un *valor propio* de T si existe un vector $\mathbf{x} \in \mathbb{R}^n$ no nulo tal que $T(\mathbf{x}) = \lambda\mathbf{x}$. Al vector $\mathbf{x}$ se le llama *vector propio* de T asociado al valor propio λ. Si consideramos la correspondencia entre operadores lineales y matrices que los representan, podemos establecer la definición de valor y vector propio para una matriz cuadrada A cualquiera como sigue: el número real λ es un valor propio de A si existe un vector $\mathbf{x} \in \mathbb{R}^n$ no nulo (vector propio asociado a λ) tal que $A\mathbf{x} = \lambda\mathbf{x}$. Bajo esta perspectiva se puede demostrar que λ es un valor propio del operador $T\colon \mathbb{R}^n \to \mathbb{R}^n$ si y sólo si lo es de *alguna* matriz A que lo representa (en particular, de la matriz que representa a T respecto de la base canónica de $\mathbb{R}^n$, que fue la que se obtuvo en la discusión correspondiente en la sección anterior). Para concretar la discusión, aquí trabajaremos con valores y vectores propios de matrices cuadradas A. Para pasar a los resultados correspondientes para operadores lineales, solamente piense en el operador $T(\mathbf{x}) = A\mathbf{x}$.

Si λ es un valor propio de la matriz A de orden n, entonces existe $\mathbf{x} \in \mathbb{R}^n$ distinto de cero tal que $A\mathbf{x} = \lambda\mathbf{x}$, expresión que se puede escribir como $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Vemos pues que el sistema homogéneo de ecuaciones lineales dado por $(A - \lambda I)\mathbf{x} = \mathbf{0}$ posee soluciones no triviales (el vector $\mathbf{x}$ es no nulo). Esto ocurre cuando (y sólo cuando) la matriz de coeficientes del sistema es no inversible, o equivalentemente cuando su determinante es igual a cero. Así pues, λ es valor propio de A si y sólo si $\det(A - \lambda I) = 0$.

Nótese que la expresión $\det(A - \lambda I)$ es un polinomio en λ de grado n, llamado *polinomio característico* de A, que denotaremos por $p(\lambda)$. Así entonces hemos establecido el siguiente resultado

Teorema 1.8.1 Los valores propios de la matriz cuadrada A son las raíces de su polinomio característico $p(\lambda) = \det(A - \lambda I)$. ∎

En toda esta obra ponemos nuestro interés en considerar *solamente raíces reales del polinomio característico* $p(\lambda)$, de modo que una matriz cuadrada A de orden n tiene, según el teorema fundamental del álgebra, *cuando mucho* n valores propios distintos, pudiéndose dar el caso de que no tenga valores propios (si todas las raíces de $p(\lambda)$ son complejas).

Ejemplo 1. Consideremos la matriz

$$A = \begin{bmatrix} 1 & 6 \\ 1 & 0 \end{bmatrix}$$

Su polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 6 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2)$$

Entonces los valores propios de A son $\lambda_1 = 3$ y $\lambda_2 = -2$. Para encontrar vectores propios asociados a estos valores propios, debemos obtener las soluciones no triviales del sistema homogéneo de ecuaciones $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Para $\lambda_1 = 3$ se tiene el sistema

$$\begin{bmatrix} -2 & 6 \\ 1 & -3 \end{bmatrix} \mathbf{x} = 0$$

de donde se obtiene que si $\mathbf{x} = (a, b)$, entonces $a = 3b$. Así, los vectores propios asociados al valor propio $\lambda_1 = 3$ son del tipo $\mathbf{x} = t(3, 1)$. De igual modo se encuentra que los vectores propios asociados al valor propio $\lambda_2 = -2$ son del tipo $\mathbf{x} = t(-2, 1)$. ■

Un problema importante en el que intervienen los valores y vectores propios de una matriz, es el problema de la *diagonalización* de la matriz. Se dice que la matriz cuadrada A de orden n es *diagonalizable* si existe una matriz inversible P de orden n de modo que $D = P^{-1}AP$ es una matriz diagonal (es decir, D es una matriz que tiene ceros fuera de su diagonal principal). Se puede demostrar que una condición equivalente a que la matriz A sea diagonalizable, es que exista una base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ del espacio $\mathbb{R}^n$ formada por vectores propios de la matriz A. De hecho, la estructura de la matriz P de la que se habló anteriormente es tal que en sus columnas están los vectores propios $\mathbf{v}_i$ de la base β.

Ejemplo 2. Considere la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Se trata de investigar si la matriz A es diagonalizable. El polinomio característico de A es

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^3$$

por lo que el único valor propio que tiene A es $\lambda = 1$. Los vectores propios correspondientes se obtienen de la solución del sistema $(A - I)\mathbf{x} = \mathbf{0}$, o sea

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

Si $\mathbf{x} = (a, b, c)$, vemos que para que $\mathbf{x}$ sea solución de este sistema (para que sea vector propio de A asociado a 1), se debe tener $b = 0$. Así, los vectores propios son del tipo $\mathbf{x} = (a, 0, c) = a(1, 0, 0) + c(0, 0, 1)$. Se ve entonces que $(1, 0, 0)$ y $(0, 0, 1)$ son dos vectores propios linealmente independientes y que no hay posibilidades de obtener otro vector propio de A linealmente independiente de estos dos vectores. Es decir, no hay manera de obtener una base de $\mathbb{R}^3$ formada por vectores propios. Así pues, concluimos que A no es diagonalizable. ■

Ejemplo 3. Consideremos la matriz

$$A = \begin{bmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{bmatrix}$$

Su polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} -2-\lambda & 8 & 6 \\ -4 & 10-\lambda & 6 \\ 4 & -8 & -4-\lambda \end{bmatrix} = -\lambda(\lambda - 2)^2$$

por lo que sus valores propios son $\lambda_1 = 2$, $\lambda_2 = 0$. Para obtener los vectores propios asociados al valor propio $\lambda_1 = 2$, resolvemos el sistema $(A - 2I)\mathbf{x} = \mathbf{0}$. Es decir,

$$\begin{bmatrix} -4 & 8 & 6 \\ -4 & 8 & 6 \\ 4 & -8 & -6 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

de donde se ve que siendo $\mathbf{x} = (a, b, c)$, éste es solución del sistema cuando $-2a + 4b + 3c = 0$. Es decir, los vectores propios asociados a $\lambda_1 = 2$ son del tipo

$$\mathbf{x} = (a, b, c) = (a, b, (2a - 4b)/3) = (a/3)(3, 0, 2) + (b/3)(0, 3, -4)$$

de donde se obtiene que los vectores $\mathbf{v}_1 = (3, 0, 2)$ y $\mathbf{v}_2 = (0, 3, -4)$ son dos vectores propios linealmente independientes asociados a este valor propio. Haciendo lo mismo con el valor propio $\lambda_2 = 0$, obtenemos el vector propio $\mathbf{v}_3 = (1, 1, -1)$. Tenemos así tres vectores propios, los cuales se puede verificar fácilmente que son l.i.[1] Entonces podemos formar la base $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ de $\mathbb{R}^3$ la cual está formada por vectores propios de A. Así esta matriz es diagonalizable. De hecho, la matriz P (tal que $P^{-1}AP$ es una matriz diagonal) tiene en sus columnas a estos vectores, es decir

$$P = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 2 & -4 & -1 \end{bmatrix}$$

La inversa de esta matriz es

$$P^{-1} = \begin{bmatrix} -1/3 & 4/3 & 1 \\ -2/3 & 5/3 & 1 \\ 2 & -4 & -3 \end{bmatrix}$$

[1] Es un hecho general que vectores propios que proceden de valores propios diferentes, son linealmente independientes.

y se tiene

$$P^{-1}AP = \begin{bmatrix} -1/3 & 4/3 & 1 \\ -2/3 & 5/3 & 1 \\ 2 & -4 & -3 \end{bmatrix} \begin{bmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 2 & -4 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

la cual es una matriz diagonal que tiene los valores propios de A en su diagonal principal (en el orden en que fueron puestos en las columnas de P los vectores propios correspondientes). ■

Un caso particular muy importante cuando se estudia el problema de la diagonalización de matrices es el caso de la diagonalización de matrices simétricas (es decir, matrices cuadradas A que coinciden con sus transpuestas). En tal caso este problema siempre tiene solución (i.e. siempre es posible diagonalizar una matriz simétrica) y además, con algunas características adicionales sobre la matriz P tal que $P^{-1}AP$ es la matriz diagonal: resulta que en este caso se puede conseguir que la matriz P sea una matriz ortogonal; es decir, que su inversa sea su transpuesta.

Presentamos a continuación un resumen de los hechos más importantes (sin demostración) que hay en torno al problema de diagonalización de matrices simétricas:

(*) Una matriz simétrica tiene todos sus valores propios reales.

(*) Los vectores propios que provienen de valores propios distintos para una matriz simétrica, son ortogonales.

(*) Siendo A una matriz simétrica, existe una matriz ortogonal[2] P tal que $P^{-1}AP = P^{t}AP$ es una matriz diagonal (se dice entonces que la matriz A es *diagonalizable ortogonalmente*).

Ejemplo 4. Consideremos la matriz simétrica

$$A = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}$$

Su polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & 4 \\ 4 & -3-\lambda \end{bmatrix} = \lambda^2 - 25$$

por lo que sus valores propios son $\lambda_1 = 5$ y $\lambda_2 = -5$. Para $\lambda_1 = 5$ se consigue el vector propio $\mathbf{x}_1 = (2, 1)$ y para $\lambda_2 = -5$ se consigue el vector propio $\mathbf{x}_2 = (1, -2)$. Obsérvese que los vectores propios $\mathbf{x}_1$ y $\mathbf{x}_2$ son ortogonales. Más aún, normalizándolos obtenemos los vectores

$$\mathbf{v}_1 = \frac{1}{\sqrt{5}}(2, 1) \qquad \mathbf{v}_2 = \frac{1}{\sqrt{5}}(1, -2)$$

[2]Las siguientes afirmaciones sobre la matriz cuadrada P de orden n son equivalentes
(1) P es ortogonal ($P^{-1} = P^{t}$)
(2) Las columnas de P forman una base ortonormal de $\mathbb{R}^n$.
(3) Las líneas de P forman una base ortonormal de $\mathbb{R}^n$.

los cuales constituyen una base ortonormal de $\mathbb{R}^2$. En este caso la matriz P es en este caso (sus columnas son los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$)

$$P = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$$

Ciertamente esta es una matriz ortogonal (se verifica fácilmente que $P^{-1} = P^t$). De hecho, se tiene

$$\begin{aligned} P^{-1}AP = P^tAP &= \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 0 \\ 0 & -5 \end{bmatrix} \end{aligned}$$

■

Ejemplo 5. Diagonalicemos ortogonalmente la matriz simétrica

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$$

Su polinomio característico es

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) = \det\begin{bmatrix} 6-\lambda & -2 & 2 \\ -2 & 5-\lambda & 0 \\ 2 & 0 & 7-\lambda \end{bmatrix} &= -\lambda^3 + 18\lambda^2 - 99\lambda + 162 \\ &= -(\lambda - 3)(\lambda - 6)(\lambda - 9) \end{aligned}$$

Tenemos entonces los tres valores propios $\lambda_1 = 3$, $\lambda_2 = 6$ y $\lambda_3 = 9$. Para el valor propio $\lambda_1 = 3$, resolvemos el sistema homogéneo $(A - 3I)\mathbf{x} = 0$ para obtener vectores propios. Eliminando coeficientes del sistema en la matriz se llega a

$$\begin{bmatrix} 3 & -2 & 2 \\ -2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

de donde se obtiene que siendo $\mathbf{x} = (a, b, c)$, se debe tener $a = b = -2c$. Así por ejemplo, $\mathbf{x}_1 = (2, 2, -1)$ es un vector propio asociado a $\lambda_1 = 3$. Para $\lambda_2 = 6$ resolvemos el sistema $(A - 6I)\mathbf{x} = 0$. Haciendo eliminación

$$\begin{bmatrix} 0 & -2 & 2 \\ -2 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

de donde se obtiene que si $\mathbf{x} = (a, b, c)$, se debe tener $b = c = -2a$. Así por ejemplo, $\mathbf{x}_2 = (1, -2, -2)$ es un vector propio asociado a $\lambda_2 = 6$. Finalmente, para $\lambda_3 = 9$ resolvemos el sistema $(A - 9I)\mathbf{x} = 0$. Eliminando

$$\begin{bmatrix} -3 & -2 & 2 \\ -2 & -4 & 0 \\ 2 & 0 & -2 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

de modo que si $\mathbf{x} = (a, b, c)$, se debe tener $a = c = -2b$. Entonces un vector propio asociado $\lambda_3 = 9$ es $\mathbf{x}_3 = (2, -1, 2)$. Se han conseguido entonces 3 vectores propios $\mathbf{x}_1 = (2, 2, -1)$, $\mathbf{x}_2 = (1, -2, -2)$, $\mathbf{x}_3 = (2, -1, 2)$ los cuales, se ve fácilmente, forman un conjunto ortogonal. Normalizándolos, obtenemos

$$\mathbf{v}_1 = (2/3, 2/3, -1/3), \qquad \mathbf{v}_2 = (1/3, -2/3, -2/3), \qquad \mathbf{v}_3 = (2/3, -1/3, 2/3)$$

que son vectores que constituyen una base ortonormal de $\mathbb{R}^3$. Entonces la matriz ortogonal P es

$$P = \begin{bmatrix} 2/3 & 1/3 & 2/3 \\ 2/3 & -2/3 & -1/3 \\ -1/3 & -2/3 & 2/3 \end{bmatrix}$$

Haciendo $P^{-1}AP$ obtenemos

$$\begin{aligned} P^{-1}AP &= P^{t}AP \\ &= \begin{bmatrix} 2/3 & 2/3 & -1/3 \\ 1/3 & -2/3 & -2/3 \\ 2/3 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 & 2/3 \\ 2/3 & -2/3 & -1/3 \\ -1/3 & -2/3 & 2/3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix} \end{aligned}$$

como tenía que ocurrir. ∎

Ejercicios (Capítulo 1, Sección 8)

En los ejercicios 1–26, determine, para cada una de las matrices dadas: **a** su polinomio característico, **b** sus valores propios, **c** los vectores propios correspondientes a cada uno de los valores propios. (Nota: en estos ejercicios todos los valores propios que resultan de las matrices, son números enteros).

1. $\begin{bmatrix} 4 & -2/5 \\ 5 & 1 \end{bmatrix}$ **2.** $\begin{bmatrix} 4 & 6 \\ 1 & 3 \end{bmatrix}$ **3.** $\begin{bmatrix} 7 & -4 \\ 2 & 1 \end{bmatrix}$

4. $\begin{bmatrix} -2 & -2 \\ 3 & 5 \end{bmatrix}$ **5.** $\begin{bmatrix} 10 & -7/5 \\ 5 & 2 \end{bmatrix}$ **6.** $\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$

7. $\begin{bmatrix} 9 & -8 \\ 1 & 3 \end{bmatrix}$ **8.** $\begin{bmatrix} -5 & -1 \\ 2 & -2 \end{bmatrix}$ **9.** $\begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}$

10. $\begin{bmatrix} 6 & 0 \\ 0 & 7 \end{bmatrix}$ **11.** $\begin{bmatrix} 1 & 1 \\ 0 & 4 \end{bmatrix}$ **12.** $\begin{bmatrix} -3 & 2 \\ 1 & -2 \end{bmatrix}$

13. $\begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & 1 \\ -2 & 2 & 1 \end{bmatrix}$ **14.** $\begin{bmatrix} 7 & 0 & -20/3 \\ 2 & -3 & 1 \\ 12/5 & -3 & 1 \end{bmatrix}$ **15.** $\begin{bmatrix} 3 & 0 & 1 \\ 1 & -4 & 1 \\ 4 & -8 & 2 \end{bmatrix}$

16. $\begin{bmatrix} 3 & 0 & -14/3 \\ 2 & 3 & 1 \\ -6 & 6 & 2 \end{bmatrix}$ **17.** $\begin{bmatrix} 4 & 0 & 27/8 \\ 4 & 4 & 4 \\ 88/27 & 4 & 4 \end{bmatrix}$ **18.** $\begin{bmatrix} 0 & 0 & -28/27 \\ 3 & 3 & 1 \\ 81/7 & 9 & 3 \end{bmatrix}$

19. $\begin{bmatrix} 4 & 0 & -5/2 \\ 2 & -3 & 1 \\ 2 & -3 & 1 \end{bmatrix}$ **20.** $\begin{bmatrix} -4 & 0 & -1/8 \\ 2 & -4 & 1 \\ 56 & 8 & -2 \end{bmatrix}$ **21.** $\begin{bmatrix} 1 & -28/3 & 0 \\ 27/14 & 2 & 3 \\ 3 & 4 & 6 \end{bmatrix}$

22. $\begin{bmatrix} 1 & -2/3 & 0 \\ -3 & 1 & 3 \\ 1 & 1 & 3 \end{bmatrix}$ **23.** $\begin{bmatrix} 1 & -2 & 0 \\ -2 & -1 & -2 \\ 2 & 1 & 2 \end{bmatrix}$ **24.** $\begin{bmatrix} 2 & -5/8 & 0 \\ 8/5 & 3 & -6/5 \\ 4 & 5 & -2 \end{bmatrix}$

25. $\begin{bmatrix} 1 & -7/2 & 0 \\ 6/7 & 2 & 6 \\ -1 & -2 & -6 \end{bmatrix}$ **26.** $\begin{bmatrix} 4 & -7/6 & 0 \\ 18/7 & 5 & -15/7 \\ 6 & 7 & -3 \end{bmatrix}$

27. Considere la matriz

$$A = \begin{bmatrix} 2 & -3 & 1 & 2 \\ 1 & -2 & 3 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Demuestre que el polinomio característico de A es $p(\lambda) = (\lambda^2 - 1)(\lambda^2 - 4)$. Obtenga entonces los valores propios de A y vectores propios correspondientes a cada uno de ellos.

28. Sea A una matriz cuadrada de orden 2. Demuestre que el polinomio característico de A es $p(\lambda) = \lambda^2 - (\text{tr } A)\lambda + \det A$, donde tr A es (la traza de A, que es igual a) la suma de los elementos de su diagonal principal.

29. Demuestre que el término independiente del polinomio característico de la matriz A, es det A.

30. Demuestre que la matriz cuadrada A es no inversible si y sólo si $\lambda = 0$ es un valor propio de A.

31. Demuestre que si λ es un valor propio de la matriz A, entonces λ^n es un valor propio de la matriz A^n, $n \in \mathbb{N}$. ¿Cómo son los vectores propios de A^n asociados al valor propio λ^n?

32. Sea A una matriz inversible y sea λ un valor propio de A (el cual es diferente de cero, según lo probado en el ejercicio 30). Demuestre que λ^{-1} es un valor propio de la matriz inversa A^{-1}. ¿Cómo son los vectores propios de A^{-1} asociados al valor propio λ^{-1}?

33. Describa los valores y vectores propios de la matriz identidad de orden n, I_n.

34. Describa los valores y vectores propios de la matriz cero de orden n.

35. Demuestre que el polinomio característico de una matriz cuadrada A es el mismo que el de su transpuesta A^t.

36. Uno de los resultados clásicos del Álgebra Lineal, es el llamado Teorema de Hamilton-Cayley: si A es una matriz cuadrada y $p(\lambda)$ es su polinomio característico, entonces $p(A) = 0$. Por ejemplo, si A es la matriz

$$A = \begin{bmatrix} 2 & -5 \\ 3 & 1 \end{bmatrix}$$

su polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 2 - \lambda & -5 \\ 3 & 1 - \lambda \end{bmatrix} = \lambda^2 - 3\lambda + 17$$

Entonces

$$p(A) = A^2 - 3A + 17I = \begin{bmatrix} 2 & -5 \\ 3 & 1 \end{bmatrix}^2 - 3\begin{bmatrix} 2 & -5 \\ 3 & 1 \end{bmatrix} + 17\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -11 & -15 \\ 9 & -14 \end{bmatrix} - \begin{bmatrix} 6 & -15 \\ 9 & 3 \end{bmatrix} + \begin{bmatrix} 17 & 0 \\ 0 & 17 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

a. Verifique el Teorema de Hamilton-Cayley para las matrices de los ejercicios 2, 3 y 13.

b. Critique la siguiente "demostración" del Teorema de Hamilton-Cayley: el polinomio característico de A es $p(\lambda) = \det(A - \lambda I)$. Entonces $p(A) = \det(A - AI) = \det(A - A) = \det \mathbf{0} = 0$.

En cada uno de los ejercicios 37–42, demuestre que la matriz dada es diagonalizable, obteniendo una base de $\mathbb{R}^2$ (o en su caso $\mathbb{R}^3$) formada por vectores propios de A. Determine la matriz P tal que $P^{-1}AP$ es una matriz diagonal.

37. La matriz del ejercicio 1.

38. La matriz del ejercicio 4.

39. La matriz del ejercicio 7.

40. La matriz del ejercicio 8.

41. La matriz del ejercicio 15.

42. La matriz del ejercicio 19.

43. Suponga que la matriz A es diagonalizable. Existe entonces una matriz inversible P tal que $P^{-1}AP = D$, donde D es una matriz diagonal. Demuestre que $P^{-1}A^nP = D^n$, $n \in \mathbb{N}$. Observando que la matriz D^n se puede obtener fácilmente elevando a la n-ésima potencia los elementos de la diagonal principal de D, concluya que la n-ésima potencia de la matriz A se puede escribir explícitamente como $A^n = PD^nP^{-1}$. Aplique estas ideas para obtener expresiones para las n-ésimas potencias de las siguientes matrices:

a. La matriz del ejercicio 2.

b. La matriz del ejercicio 3.

c. La matriz del ejercicio 8.

44. Considere la matriz del ejemplo 3

$$A = \begin{bmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{bmatrix}$$

Demuestre que la n-ésima potencia de esta matriz es

$$A^n = 2^n \begin{bmatrix} -1 & 4 & 3 \\ -2 & 5 & 3 \\ 2 & -4 & -2 \end{bmatrix}$$

45. Una *sucesión de Fibonacci* es una sucesión de números enteros $(a_0, a_1, a_2, \ldots)$ en la que $a_n = a_{n-1} + a_{n-2}$, $n = 2, 3, \ldots$, y a_0 y a_1 son números dados. La fórmula $a_n = a_{n-1} + a_{n-2}$ que define el término a_n de la sucesión de Fibonacci se dice ser una *fórmula de recurrencia*, pues a_n se determina una vez determinados sus dos términos anteriores. El objetivo de este ejercicio es obtener una fórmula *explícita* para el término a_n de una sucesión de Fibonacci. Considere el caso en el que $a_0 = 0$, $a_1 = 1$. Así, la sucesión de Fibonacci se ve como $(0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots)$.

a. Considere la matriz $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Para $n = 0, 1, 2, \ldots$ defina el vector $\mathbf{x}_n \in \mathbb{R}^2$ como $\mathbf{x}_n = A^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Use inducción para probar que $\mathbf{x}_n = (a_{n+1}, a_n)$, donde (a_n) es la sucesión de Fibonacci que se está considerando.

b. Hallando una fórmula explícita para A^n, demuestre que

$$a_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right]$$

(¡esta fórmula define un número *entero*!).

c. Demuestre que $\lim \frac{a_{n+1}}{a_n} = \frac{1+\sqrt{5}}{2}$. Los griegos llamaban a este número "razón áurea".

En los ejercicios 46–53, diagonalice ortogonalmente la matriz simétrica dada, obteniendo una matriz ortogonal P tal que $P^{-1}AP = P^{t}AP$ sea una matriz diagonal.

46. $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ **47.** $\begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}$ **48.** $\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$

49. $\begin{bmatrix} 8 & -5 \\ -5 & 8 \end{bmatrix}$ **50.** $\begin{bmatrix} -4 & 2 \\ 2 & -4 \end{bmatrix}$ **51.** $\begin{bmatrix} 0 & 5 \\ 5 & 0 \end{bmatrix}$

52. $\begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$ **53.** $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$

1.9 Formas cuadráticas

En esta sección veremos algunos hechos importantes que aparecen en el estudio de las formas cuadráticas. El material que aquí se presenta se usará en el estudio de los extremos de las funciones de varias variables en el capítulo 4.

Una *forma cuadrática* en $\mathbb{R}^n$ es una función $q: \mathbb{R}^n \to \mathbb{R}$ definida por

$$q(\mathbf{x}) = [\mathbf{x}]^t A[\mathbf{x}]$$

donde A es una matriz simétrica de orden n y $[\mathbf{x}]$ es la matriz de coordenadas de $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ (entonces $[\mathbf{x}]$ es una matriz $1 \times n$, y $[\mathbf{x}]^t$, la transpuesta de $[\mathbf{x}]$, es una matriz $n \times 1$).

Por ejemplo, una forma cuadrática en $\mathbb{R}^2$ es una función $q: \mathbb{R}^2 \to \mathbb{R}$

$$q(x, y) = [x\ y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + cy^2 + 2bxy$$

y una forma cuadrática en $\mathbb{R}^3$ es del tipo $q: \mathbb{R}^3 \to \mathbb{R}$

$$q(x, y, z) = [x\ y\ z] \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
$$= ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz$$

En general, una forma cuadrática en $\mathbb{R}^n$ es una función de la forma

$$q(x_1, x_2, \ldots, x_n) = [x_1\ x_2\ \ldots\ x_n] \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} a_{ii}x_i^2 + 2 \sum_{\substack{i,j=1 \\ i<j}}^{n} a_{ij}x_ix_j$$

Ejemplo 1. Consideremos la forma cuadrática $q: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$q(x, y, z) = 7x^2 + 5y^2 - 4xy + 26xz + 30yz$$

Según la fórmula recientemente obtenida para una forma cuadrática general, podemos escribir a q como $q(x, y, z) = [x\ y\ z]A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ donde $A = (a_{ij})$ es la matriz simétrica ($a_{ij} = a_{ji}$, $i, j = 1, 2, 3$) para la cual $a_{11} = 7$ (coeficiente de $x_1 = x$), $a_{22} = 5$ (coeficiente de $x_2 = y$), $a_{33} = 0$ (coeficiente de $x_3 = z$), $2a_{12} = -4$ (coeficiente de $x_1x_2 = xy$), $2a_{13} = 26$ (coeficiente de $x_1x_3 = xz$), $2a_{23} = 30$ (coeficiente de $x_2x_3 = yz$). Entonces

$$A = \begin{bmatrix} 7 & -2 & 13 \\ -2 & 5 & 15 \\ 13 & 15 & 0 \end{bmatrix}$$ ■

Una forma cuadrática $q: \mathbb{R}^n \to \mathbb{R}$ se dice ser:

a. *definida positiva*, si $q(\mathbf{x}) > 0\ \forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

b. *definida negativa*, si $q(\mathbf{x}) < 0\ \forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

c. *semidefinida positiva*, si $q(\mathbf{x}) \geq 0\ \forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

d. *semidefinida negativa*, si $q(\mathbf{x}) \leq 0\ \forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

Ejemplo 2. La forma cuadrática $q: \mathbb{R} \to \mathbb{R}$, dada por $q(x) = ax^2$ es definida positiva (definida negativa, semidefinida positiva, semidefinida negativa) si y sólo si $a > 0$ ($a < 0$, $a \geq 0$, $a \leq 0$, respectivamente). ■

Ejemplo 3. La forma cuadrática $q: \mathbb{R}^2 \to \mathbb{R}$ dada por $q(x, y) = 3x^2 + 7y^2$ es definida positiva, pues es claro que $q(x, y) > 0\ \forall (x, y) \in \mathbb{R}^2$, $(x, y) \neq (0, 0)$. Del mismo modo se ve claro que la forma cuadrática $q: \mathbb{R}^3 \to \mathbb{R}$ dada por $q(x, y, z) = -x^2 - 8y^2 - 10z^2$ es definida negativa y que la forma cuadrática $q: \mathbb{R}^3 \to \mathbb{R}$, $q(x, y, z) = 2x^2 + 3z^2$ es semidefinida positiva, pues en este caso $q(x, y, z) \geq 0\ \forall (x, y, z) \neq (0, 0, 0)$ (de hecho, obsérvese que para todos los vectores del tipo $(0, y, 0)$ la forma q vale 0). ■

Con los dos ejemplos anteriores podemos darnos cuenta de que la propiedad de una forma cuadrática de ser definida positiva o negativa se descrubre inmediatamente cuando ésta solamente

posee los términos cuadráticos. Más en concreto, podemos decir que la forma cuadrática $q\colon \mathbb{R}^n \to \mathbb{R}$ dada por

$$q(x_1, x_2, \ldots, x_n) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

será definida positiva (definida negativa, semidefinida positiva, semidefinida negativa) si y sólo si todos los coeficientes λ_i son números positivos (negativos, no negativos, no positivos, respectivamente). Nótese que en este caso la matriz simétrica A de la forma cuadrática es en realidad una matriz *diagonal* (cuyos elementos en la diagonal principal son los números λ_i).

Ejemplo 4. Consideremos la forma cuadrática $q\colon \mathbb{R}^3 \to \mathbb{R}$ dada por

$$q(x, y, z) = 6x^2 + 5y^2 + 7z^2 - 4xy + 4xz$$

o bien

$$q(x, y, z) = [x\ y\ z]A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

en la que

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$$

Ciertamente no es posible afirmar "a golpe de vista" que esta forma cuadrática es definida positiva (veremos a continuación que de hecho lo es). Sin embargo, si escribiéramos nuestra forma cuadrática como

$$q(x, y, z) = 3(2x/3 + 2y/3 - z/3)^2 + 6(x/3 - 2y/3 - 2z/3)^2 + 9(2x/3 - y/3 + 2z/3)^2$$

(queda de responsabilidad del lector verificar que esta expresión es *efectivamente* igual a la expresión de $q(x, y, z)$ que se dio al inicio del problema), resultaría evidente que la forma es definida positiva, pues es una suma de cuadrados con coeficientes positivos. ■

El ejemplo anterior nos muestra que la propiedad de que una forma cuadrática sea definida positiva o negativa, se descubre fácilmente si escribimos la expresión que define a la forma de una "manera adecuada"; por ejemplo, si la escribimos como una suma de cuadrados. Se nos plantean así dos preguntas naturales: 1) ¿Una forma cuadrática se puede escribir siempre como una suma de cuadrados?, 2) De ser así, ¿cómo se procede para lograrlo? La respuesta a la pregunta 1) es —afortunadamente— afirmativa. Al tratar de ver el por qué de esto, se dará respuesta a la pregunta 2).

Tomemos entonces la forma cuadrática $q\colon \mathbb{R}^n \to \mathbb{R}$, $q(\mathbf{x}) = [\mathbf{x}]^t A[\mathbf{x}]$, en donde A es una matriz simétrica. En la sección anterior se veia que A por ser simétrica, es diagonalizable ortogonalmente; es decir, que existe una matriz ortogonal P tal que $P^t AP$ es una matriz diagonal. Se vio también que las columnas de P forman una base ortonormal de $\mathbb{R}^n$. Sea $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ dicha base. Entonces los vectores $\mathbf{v}_i$ son vectores propios de A, los cuales tienen norma 1 y son ortogonales dos a dos. Escribamos el vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ en términos de la base β; digamos que $[\mathbf{x}]_\beta = (x'_1, x'_2, \ldots, x'_n)$, o bien, como matriz

$$[\mathbf{x}]_\beta = \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}$$

Observemos entonces que, puesto que $\mathbf{x} = x'_1\mathbf{v}_1 + x'_2\mathbf{v}_2 + \cdots + x'_n\mathbf{v}_n$, podemos escribir

$$[\mathbf{x}] = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}$$

Entonces, la forma $q(x)$ queda como (recordando que $(AB)^t = B^t A^t$)

$$\begin{aligned} q(\mathbf{x}) &= [\mathbf{x}]^t A[\mathbf{x}] = (P[\mathbf{x}]_\beta)^t A(P[\mathbf{x}]_\beta) \\ &= [\mathbf{x}]_\beta^t (P^t A P)[\mathbf{x}]_\beta = [\mathbf{x}]_\beta^t D[\mathbf{x}]_\beta \end{aligned}$$

donde D es una matriz diagonal la cual, sabemos, tiene en su diagonal principal los valores propios de A (que aparecen en el mismo orden en que fueron acomodados los vectores propios correspondientes en las columnas de P). Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales valores propios. Escribiendo explícitamente la última expresión nos queda que

$$\begin{aligned} q(x) &= [x]_\beta^t D[x]_\beta \\ &= [x'_1 \; \ldots \; x'_n] \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} \\ &= \lambda_1(x'_1)^2 + \lambda_2(x'_2)^2 + \cdots + \lambda_n(x'_n)^2 \end{aligned}$$

Así pues, hemos logrado escribir a $q(x)$ como una suma de cuadrados de las coordenadas x'_i del vector $\mathbf{x}$ en términos de la base ortonormal β. No es difícil ver que tales coordenadas se pueden calcular como productos internos del vector $\mathbf{x}$ con los vectores de la base β (esto ocurre porque β es una base ortonormal). Es decir que $x'_i = \mathbf{x} \cdot \mathbf{v}_i$.

En resumen tenemos que: la forma cuadrática $q: \mathbb{R}^n \to \mathbb{R}$, $q(x) = [\mathbf{x}]^t A[\mathbf{x}]$, donde A es una matriz simétrica, se puede escribir como

$$q(\mathbf{x}) = \lambda_1(\mathbf{x} \cdot \mathbf{v}_1)^2 + \lambda_2(\mathbf{x} \cdot \mathbf{v}_2)^2 + \cdots + \lambda_n(\mathbf{x} \cdot \mathbf{v}_n)^2$$

donde λ_i son los valores propios de A, y $\mathbf{v}_i$ son vectores propios (normalizados) correspondientes a tales valores propios.

Regresando al ejemplo 4 anterior, vemos que la matriz

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$$

tiene por valores propios a $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$ (ver ejemplo 5 de la sección anterior) y que los vectores

$$\mathbf{v}_1 = (2/3, 2/3, -1/3), \qquad \mathbf{v}_2 = (1/3, -2/3, -2/3), \qquad \mathbf{v}_3 = (2/3, -1/3, 2/3)$$

de donde

$$\begin{aligned} q(x, y, z) = q(u) &= \lambda_1(u \cdot v_1)^2 + \lambda_2(u \cdot v_2)^2 + \lambda_3(u \cdot v_3)^2 \\ &= 3(2x/3 + 2y/3 - z/3)^2 + 6(x/3 - 2y/3 - 2z/3)^2 + 9(2x/3 - y/3 + 2z/3)^2 \end{aligned}$$

que es el resultado que se usó en el ejemplo mencionado.

De la discusión anterior se desprende inmediatamente el siguiente teorema

Teorema 1.9.1 La forma cuadrática $q: \mathbb{R}^n \to \mathbb{R}$, $q(\mathbf{x}) = [\mathbf{x}]^t A[\mathbf{x}]$, en donde A es una matriz simétrica, es definida positiva (definida negativa, semidefinida positiva, semidefinida negativa) si y sólo si todos los valores propios de A son positivos (negativos, no negativos, no positivos, respectivamente). ■

Ejemplo 5. Consideremos la forma cuadrática $q: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$q(x, y, z) = [x\ y\ z] \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Para investigar si esta forma cuadrática es definida positiva, definida negativa, etc. sólo tenemos que ver los valores propios de la matriz simétrica que define a la forma q. El polinomio característico de esta matriz es

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 2 & -2 \\ 2 & 5-\lambda & -4 \\ -2 & -4 & 5-\lambda \end{bmatrix} = -(\lambda - 1)^2(\lambda - 10)$$

Así pues, los valores propios de A son $\lambda_1 = 10$, $\lambda_2 = 1$. Como éstos son positivos, concluimos que la forma dada es definida positiva. ■

Diremos que una matriz simétrica A de orden n es definida positiva, definida negativa, semidefinida positiva o semidefinida negativa, si la forma cuadrática correspondiente $q: \mathbb{R}^n \to \mathbb{R}$, $q(x) = [\mathbf{x}]^t A[\mathbf{x}]$ lo es.

Ejemplo 6. Consideremos la matriz simétrica

$$A = \begin{bmatrix} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & 5 \end{bmatrix}$$

El polinomio característico de A es

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) &= \det \begin{bmatrix} 11-\lambda & 2 & -8 \\ 2 & 2-\lambda & 10 \\ -8 & 10 & 5-\lambda \end{bmatrix} \\ &= -(\lambda - 9)(\lambda + 9)(\lambda - 18) \end{aligned}$$

Los valores propios de A son entonces $\lambda_1 = 9$, $\lambda_2 = -9$, $\lambda_3 = 18$. En este caso concluimos que la matriz no cae dentro de ninguna de las clasificaciones estudiadas aquí. De hecho, si consideramos la correspondiente forma cuadrática $q: \mathbb{R}^3 \to \mathbb{R}$,

$$q(x, y, z) = [x\ y\ z]A[x\ y\ z]^t$$

se tiene que los valores que toma esta forma son positivos para algunos vectores de $\mathbb{R}^3$ y negativos para otros (por ejemplo, $q(0, -1, 1) = -13 < 0$, pero $q(1, 1, 0) = 17 > 0$). ∎

A una matriz (o a una forma cuadrática) como la del ejemplo anterior, se le llama *indefinida*.

En el siguiente teorema se establece otro criterio útil para decidir si una matriz simétrica A es definida positiva o definida negativa. La demostración de este resultado puede consultarse en [*Pi*], pp. 624 a 627. En este teorema se usa el concepto de *submatriz angular*: consideremos la matriz cuadrada $A = (a_{ij})$ de orden n. A las submatrices

$$A_1 = [a_{11}],\quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\quad A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \ldots, A_n = A$$

se les llama "submatrices angulares de A". En forma esquemática

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Teorema 1.9.2 Sea $q: \mathbb{R}^n \to \mathbb{R}$ la forma cuadrática $q(\mathbf{x}) = [\mathbf{x}]^t A[\mathbf{x}]$. Esta forma (o bien, la matriz A) es definida positiva (respectivamente, definida negativa) si y sólo si los determinantes de las submatrices angulares $\Delta_1 = \det A_1$, $\Delta_2 = \det A_2, \ldots, \Delta_n = \det A_n = \det A$, son positivos (los determinantes tienen signos alternados, comenzando con $\Delta_1 = \det A_1 < 0$, respectivamente). ∎

NOTA: El resultado anterior sigue siendo válido si aceptamos el cero en los determinantes de las submatrices angulares, cambiando entonces la conclusión correspondiente a "formas semidefinidas positivas" o bien, "formas semidefinidas negativas". Es claro que si los determinantes de las submatrices angulares no siguen alguno de los patrones establecidos en este resultado, la forma es indefinida.

Ejemplo 7. Consideremos la matriz simétrica del ejemplo 5

$$A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$$

Los determinantes de las submatrices angulares de A son

$$\Delta_1 = \det A_1 = \det[2] = 2 > 0$$
$$\Delta_2 = \det A_2 = \det \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} = 6 > 0$$
$$\Delta_3 = \det A_3 = \det A = \det \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} = 10 > 0$$

Puesto que todos estos determinantes son positivos, concluimos que la matriz es definida positiva (como lo hicimos en el ejemplo 5 usando los valores propios de la matriz). ■

Ejemplo 8. Sea A la matriz

$$A = \begin{bmatrix} -5 & 2 & 2 \\ 2 & -6 & 0 \\ 2 & 0 & -4 \end{bmatrix}$$

Los determinantes de las submatrices angulares de A son

$$\Delta_1 = \det A_1 = \det[-5] = -5 < 0$$
$$\Delta_2 = \det A_2 = \det \begin{bmatrix} -5 & 2 \\ 2 & -6 \end{bmatrix} = 26 > 0$$
$$\Delta_3 = \det A_3 = \det A = \det \begin{bmatrix} -5 & 2 & 2 \\ 2 & -6 & 0 \\ 2 & 0 & -4 \end{bmatrix} = -80 < 0$$

Puesto que los signos de estos determinantes son alternados, comenzando con $\Delta_1 < 0$, concluimos que la matriz A es definida negativa. ■

Ejemplo 9. Consideremos la matriz del ejemplo 6

$$A = \begin{bmatrix} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & 5 \end{bmatrix}$$

Los determinantes de las submatrices angulares de A son

$$\Delta_1 = \det[11] = 11 > 0$$
$$\Delta_2 = \det \begin{bmatrix} 11 & 2 \\ 2 & 2 \end{bmatrix} = 18 > 0$$
$$\Delta_3 = \det \begin{bmatrix} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & 5 \end{bmatrix} = -1458 < 0$$

Los signos de estos determinantes no siguen alguno de los patrones comentados en el teorema 1.9.2. Concluimos entonces que la matriz A es indefinida. ■

Ejercicios (Capítulo 1, Sección 9)

En los ejercicios 1–3, escriba la forma cuadrática $q: \mathbb{R}^2 \to \mathbb{R}$ dada, como $q(x, y) = [x\ y]A \begin{bmatrix} x \\ y \end{bmatrix}$, (con A una matriz simétrica de orden 2), en la forma $q(x, y) = k_1x^2 + k_2xy + k_3y^2$, donde k_1, k_2 y k_3 son números reales.

1. $q(x, y) = [x\ y] \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

2. $q(x, y) = [x\ y] \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

3. $q(x, y) = [x\ y] \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

En los ejercicios 4–7, escriba la forma cuadrática $q: \mathbb{R}^3 \to \mathbb{R}$ dada, como $q(x, y, z) = [x\ y\ z]A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, (con A una matriz simétrica de orden 3), en la forma $q(x, y, z) = k_1x^2+k_2y^2+k_3z^2+k_4xy+k_5xz+k_6yz$

4. $q(x, y, z) = [x\ y\ z] \begin{bmatrix} 2 & 1 & 4 \\ 1 & 5 & 3 \\ 4 & 3 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

5. $q(x, y, z) = [x\ y\ z] \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

6. $q(x, y, z) = [x\ y\ z] \begin{bmatrix} 3 & 0 & 3 \\ 0 & 5 & 3 \\ 3 & 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

7. $q(x, y, z) = [x\ y\ z] \begin{bmatrix} 7 & -1 & -3 \\ -1 & -5 & 0 \\ -3 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

8. Escriba la forma cuadrática $q: \mathbb{R}^4 \to \mathbb{R}$ dada, como suma de términos del tipo $k_{ij}x_ix_j$, $i, j = 1, 2, 3, 4$.

$$q(x_1, x_2, x_3, x_4) = [x_1\ x_2\ x_3\ x_4] \begin{bmatrix} 3 & 2 & 0 & -1 \\ 2 & 5 & 4 & -2 \\ 0 & 4 & 7 & -7 \\ -1 & -2 & -7 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

En los ejercicios 9–15, escriba la forma cuadrática dada como $q(x) = [\mathbf{x}]^t A[\mathbf{x}]$, donde A es una matriz simétrica.

9. $q: \mathbb{R}^2 \to \mathbb{R}$, $q(x, y) = 3x^2 + y^2 + 4xy$

10. $q: \mathbb{R}^2 \to \mathbb{R}$, $q(x, y) = x^2 + 10y^2$

11. $q: \mathbb{R}^3 \to \mathbb{R}$, $q(x, y, z) = x^2 + y^2 + z^2 + xy + xz + yz$

12. $q: \mathbb{R}^3 \to \mathbb{R}$, $q(x, y, z) = x^2 + yz$

13. $q: \mathbb{R}^3 \to \mathbb{R}$, $q(x, y, z) = 2xy + 4xz - 10yz$

14. $q: \mathbb{R}^4 \to \mathbb{R}$, $q(x_1, x_2, x_3, x_4) = x_1x_2 + x_3x_4$

15. $q: \mathbb{R}^4 \to \mathbb{R}$, $q(x_1, x_2, x_3, x_4) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_4^2 - x_1x_3 - x_2x_4$

En los ejercicios 16–50, aplique el criterio establecido en el teorema 1.9.2 para decidir si la matriz simétrica dada (o bien, la forma cuadrática dada) es definida positiva, definida negativa, semidefinida positiva, semidefinida negativa, o indefinida.

16. $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$,

17. $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$

18. $\begin{bmatrix} -1 & 3 \\ 2 & 4 \end{bmatrix}$

19. $\begin{bmatrix} -3 & 5 \\ -2 & 1 \end{bmatrix}$

20. $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 4 & 3 \\ 1 & 3 & 11 \end{bmatrix}$

21. $\begin{bmatrix} -2 & 2 & 3 \\ 2 & -8 & 0 \\ 3 & 0 & -24 \end{bmatrix}$

22. $\begin{bmatrix} 2 & 4 & 1 \\ 4 & 16 & 3 \\ 1 & 3 & 9/16 \end{bmatrix}$

23. $\begin{bmatrix} -2 & -1 & 0 \\ -1 & -5/2 & -2 \\ 0 & -2 & -4 \end{bmatrix}$

24. $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 16 \end{bmatrix}$

25. $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

26. $\begin{bmatrix} -5 & 1 & 2 \\ 1 & -4/5 & -8 \\ 2 & -8 & -80 \end{bmatrix}$

27. $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

28. $\begin{bmatrix} -3 & 1 & 0 \\ 1 & -3 & -2 \\ 0 & -2 & -4/3 \end{bmatrix}$

29. $\begin{bmatrix} -2 & 2 & 1 \\ 2 & -8 & 0 \\ 1 & 0 & -4 \end{bmatrix}$

30. $\begin{bmatrix} -1 & 2 & 2 \\ 2 & -8 & 0 \\ 2 & 0 & -16 \end{bmatrix}$

31. $\begin{bmatrix} 4 & 1 & 0 \\ 1 & 1/2 & 3 \\ 0 & 3 & 72 \end{bmatrix}$

32. $\begin{bmatrix} 3 & 2 & -1 \\ 6 & 1 & 4 \\ 0 & 0 & 2 \end{bmatrix}$

33. $\begin{bmatrix} -1 & 3 & -2 \\ 3 & -18 & 0 \\ -2 & 0 & 0 \end{bmatrix}$

34. $\begin{bmatrix} -3 & 2 & 1 \\ 2 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}$

35. $\begin{bmatrix} -3 & 3 & 0 \\ 3 & -15 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

36. $q(x, y, z) = -3x^2 - 128y^2 - 8z^2 - 16xy - 6yz$

37. $q(x, y, z) = x^2 + 2y^2 + 8z^2 + 2xy + 4yz$

38. $q(x, y, z) = 4x^2 - y^2 - 4z^2 + 2xy + 6xz - 4yz$

39. $q(x, y, z) = -x^2 - 36y^2 - 2z^2 + 6xy + 2yz$

40. $q(x, y, z) = 8x^2 + 18y^2 + 8z^2/9 + 8xy - 2xz + 8yz$

41. $q(x, y, z) = -3x^2 - 4y^2/3 - 6z^2 + 2xy + 16yz$

42. $q(x, y, z) = -4x^2 - 16y^2 - z^2 + 8xy - 6yz$

43. $q(x, y, z) = x^2 + 8y^2 + z^2/2 + 4xy + 2yz$

44. $q(x, y, z) = x^2 + y^2 + 10z^2 + 4xz + 2yz$

45. $q(x, y, z) = -4x^2 - 4y^2 - 8z^2/3 + 4xy + 4xz$

46. $q(x, y, z) = -2x^2 + y^2 + z^2$

47. $q(x, y, z) = -5x^2 - 4y^2 - 2z^2 + 4xy - 2yz$

48. $q(x, y, z) = 4x^2 + 6y^2 + 14z^2 - 4xz - 12yz$

49. $q(x, y, z) = -2x^2 - 3y^2 - 4z^2 + 2xy + 4yz$

50. $q(x, y, z) = -2x^2 - 10y^2 - 4z^2 - 4xy + 5yz$

En los ejercicios 51–85, use el teorema 1.9.1 para decidir si la forma cuadrática dada es definida positiva, definida negativa, semidefinida positiva, semidefinida negativa o indefinida. En cada caso, escriba la forma como una suma de cuadrados. (NOTA: todos los valores propios de las matrices simétricas correspondientes, son números enteros).

51. $q(x, y) = \frac{1}{5}(34x^2 + 41y^2 + 24xy)$

52. $q(x, y) = \frac{1}{5}(61x^2 + 89y^2 + 96xy)$

53. $q(x, y) = \frac{1}{13}(469x^2 + 1897y^2 + 1440xy)$

54. $q(x, y) = \frac{1}{13}(194x^2 + 313y^2 + 120xy)$

55. $q(x, y) = 61x^2 + 89y^2 + 96xy$

56. $q(x, y) = -\frac{1}{5}(151x^2 + 249y^2 + 336xy)$

57. $q(x, y) = -\frac{1}{13}(219x^2 + 457y^2 + 240xy)$

58. $q(x, y) = \frac{1}{169}(244x^2 + 601y^2 + 360xy)$

59. $q(x, y) = \frac{1}{13}(119x^2 - 119y^2 - 240xy)$

60. $q(x, y) = -2(34x^2 + 41y^2 + 24xy)$

61. $q(x, y, z) = 5x^2 + 6y^2 + 7z^2 + 4xy + 4yz$

62. $q(x, y, z) = -18x^2 - 15y^2 - 21z^2 - 12xy + 12xz$

63. $q(x, y, z) = 25x^2 + 39y^2 + 13z^2 + 24xy - 24yz$

64. $q(x, y, z) = \frac{1}{7}(103x^2 + 125y^2 + 66z^2 + 48xy + 12xz - 60yz)$

65. $q(x, y, z) = \frac{59}{5}x^2 + \frac{66}{5}y^2 + 5z^2 + \frac{24}{5}xy$

66. $q(x, y, z) = \frac{1}{3}(203x^2 + 86y^2 + 35z^2 - 152xy + 32xz - 40yz)$

67. $q(x, y, z) = 34x^2 + 41y^2 + 5z^2 + 24xy$

68. $q(x, y, z) = 8x^2 + 9y^2 + 10z^2 + 4xy + 4yz$

69. $q(x, y, z) = \frac{1}{49}(103x^2 + 125y^2 + 66z^2 + 48xy + 12xz - 60yz)$

70. $q(x, y, z) = \frac{1}{49}(-382x^2 - 541y^2 - 155z^2 - 324xy - 60xz + 384yz)$

71. $q(x, y, z) = -10x^2 - 12y^2 - 14z^2 - 8xy - 8yz$

72. $q(x, y, z) = 6x^2 + 12y^2 + 9z^2 + 12xz + 12yz$

73. $q(x, y, z) = -\frac{84}{5}x^2 - \frac{91}{5}y^2 - 5z^2 - \frac{24}{5}xy$

74. $q(x, y, z) = \frac{1}{13}(345x^2 + 466y^2 + 203z^2 - 120xy - 48xz - 168yz)$

75. $q(x, y, z) = \frac{1}{169}(1197x^2 + 1439y^2 + 913z^2 - 240xy - 96xz - 336yz)$

76. $q(x, y, z) = \frac{1}{5}(102x^2 + 95y^2 + 73z^2 - 100xy + 28xz)$

77. $q(x, y, z) = \frac{1}{3}(433x^2 + 259y^2 + 118z^2 - 208xy + 16xz - 152yz)$

78. $q(x, y, z) = \frac{1}{27}(203x^2 + 86y^2 + 35z^2 - 152xy + 32xz - 40yz)$

79. $q(x, y, z) = \frac{59}{25}x^2 + \frac{66}{25}y^2 + z^2 + \frac{24}{25}xy$

80. $q(x, y, z) = \frac{1}{3}(-181x^2 - 10y^2 + 29z^2 + 232xy - 62xz + 8yz)$

81. $q(x, y, z) = \frac{1}{13}(192x^2 + 441y^2 + 43z^2 - 216xy - 24xz - 240yz)$

82. $q(x, y, z) = \frac{1}{13}(-649x^2 - 1915y^2 - 309z^2 + 1008xy - 96xz + 912yz)$

83. $q(x, y, z) = 33x^2 + 114y^2 + 96z^2 + 120xy - 48xz - 48yz$

84. $q(x, y, z) = \frac{422}{45}x^2 + \frac{55}{9}y^2 - \frac{247}{45}z^2 - \frac{388}{9}xy + \frac{508}{45}xz - \frac{16}{9}yz$

85. $q(x, y, z) = 7x^2 - 33y^2 + 5z^2 - 48xy + 24xz + 24yz$

Capítulo 2

Funciones de varias variables

En este capítulo trabajaremos con funciones cuyos valores, que serán números reales, dependerán de más de una variable. Por ejemplo, el área de un triángulo (que es un valor real) depende (es función de) de su base b y su altura h. Así, si T es el conjunto de todos los triángulos, podemos formar la función $f: T \to \mathbb{R}$ que asocia a cada triángulo T, su área. De otro modo, podemos acomodar en parejas ordenadas los valores de la base y altura de cada triángulo en T, de tal manera que la función f tomará la pareja $(b, h) \in \mathbb{R} \times \mathbb{R}$ y le asociará el número real $f(b, h) = \frac{1}{2}bh =$ área del triángulo de base b y altura h. Esto lo escribimos como $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $f(b, h) = \frac{1}{2}bh$. En este caso la función f solamente acepta valores positivos de b y h. Decimos entonces que el *dominio* U de f, que es un subconjunto de $\mathbb{R}^2$, estará formado por aquellas parejas (x, y) cuyas dos coordenadas sean positivas. Es decir

$$U = \mathbb{R}^+ \times \mathbb{R}^+ = \{(x, y) \in \mathbb{R}^2 | x > 0, y > 0\} \subset \mathbb{R}^2$$

En física se encuentran también muchos ejemplos de funciones reales que dependen de más de una variable. El volumen V de un globo inflado depende de la presión P del aire que contiene y de la temperatura T. Así, el número real V es función de P y T. De hecho $V: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $V(P, T) = kT/P =$ volumen del globo, donde k es una constante (ley general de los gases ideales).

Dividiremos en 3 partes el estudio del cálculo diferencial de estas funciones. La primera de ellas (de la cual se ocupa el presente capítulo) tiene por objeto estudiar el concepto de diferenciabilidad de estas funciones y algunos tópicos simples relacionados con este concepto. En la segunda parte (el capítulo 3) se estudiarán temas más específicos sobre funciones diferenciables de varias variables, como son la derivación de funciones compuestas, las funciones inversas y las funciones implícitas. Por último, en la tercera parte (el capítulo 4) se estudiará lo relacionado con los extremos de estas funciones. El estudio del cálculo integral de este tipo de funciones se hará en el capítulo 6.

2.1 Funciones de varias variables

Recordemos que una función $f: A \to B$ del conjunto A al conjunto B (ambos no vacíos) es una regla que asocia a cada elemento $a \in A$ un elemento (y sólo uno) bien determinado $b \in B$, llamado "imagen de abajo f" y escrito como $b = f(a)$, (b es igual a f de a). El conjunto A es el dominio de f, el conjunto B es su contradominio (o codominio), y el conjunto formado por todas las imágenes

de f es su rango, es decir,

$$\text{rango de } f = \{b \in B | b = f(a),\ a \in A\}$$

En el primer curso de cálculo se trabajó con funciones en las cuales A y B eran (subconjuntos de) el conjunto de números reales $\mathbb{R}$, que se escribían como $f\colon I \subseteq \mathbb{R} \to \mathbb{R}$, notación en la que se enfatiza que el dominio de la función es el conjunto I de $\mathbb{R}$ y que su codominio es $\mathbb{R}$ (el rango de la función no queda explícito en esta notación). Decíamos entonces que f es una *función real* (las imágenes $f(x)$ son números reales) *de una variable real* $x \in I$. Ahora vamos a considerar funciones cuyas imágenes son también números reales, pero cuyo dominio será un subconjunto del espacio $\mathbb{R}^n$. Es decir, funciones del tipo $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, llamadas *funciones reales de n variables reales* (viendo a los punto de $\mathbb{R}^n$ como n-adas de números reales —las variables de la función), o bien, *funciones reales de variable vectorial* (viendo a los elementos de $\mathbb{R}^n$ como vectores).

Tenemos pues que una función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una regla que asocia a cada n-ada ordenada de números reales $(x_1, x_2, \ldots, x_n)$ de U, o bien, a cada vector $\mathbf{x}$ de U, un número real bien determinado $f(x_1, x_2, \ldots, x_n)$. El conjunto U es el dominio de la función f, su codominio es $\mathbb{R}$ y el rango de f es el conjunto de $z \in \mathbb{R}$ para las cuales existe $\mathbf{x} \in U$ tal que $z = f(\mathbf{x})$, es decir

$$\text{rango de } f = \{z \in \mathbb{R} | z = f(\mathbf{x}),\ \mathbf{x} \in U\}$$

Obsérvese que éste es un subconjunto de $\mathbb{R}$. Así, cada una de estas funciones está constituida por: 1) su dominio $U \subseteq \mathbb{R}^n$, 2) su codominio $\mathbb{R}$, y, 3) la regla que asocia a cada $\mathbf{x} \in U$, el número $f(\mathbf{x}) \in \mathbb{R}$, imagen de x bajo f. Esquemáticamente se tiene

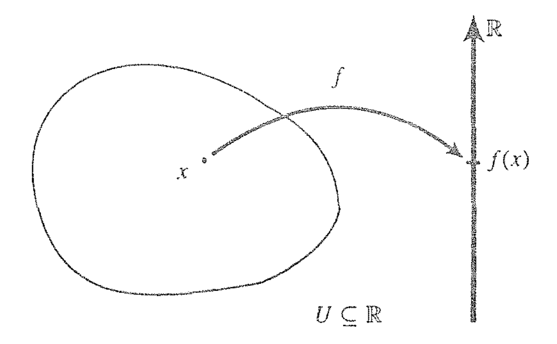


Figura 1. Una función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$.

Usaremos indistintamente las notaciones $f(x_1, x_2, \ldots, x_n)$ o $f(\mathbf{x})$ para denotar la imagen bajo f del vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Es usual, en una función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, dar simplemente la regla $z = f(\mathbf{x})$ por medio de la cual se asocia a cada vector $\mathbf{x} \in U$, el número real $z = f(x)$, y no dar explícitamente el dominio U de f (se entiende que el codominio siempre es $\mathbb{R}$). En tal caso se debe entender que el dominio de f es el mayor subconjunto U del espacio $\mathbb{R}^n$ para el cual la regla $f(x)$ tenga sentido con $\mathbf{x} \in U$ (llamado *dominio natural* de la función f).

Ejemplo 1. La función $f(x_1, x_2) = x_1^2 + x_2^2$, asocia a cada pareja $(x_1, x_2) \in \mathbb{R}^2$ el número real $x_1^2 + x_2^2$. Así, se tiene que $f(1, 2) = (1)^2 + (2)^2 = 5 =$ imagen bajo f de $(1, 2)$, $f(-3, 4) = (-3)^2 + (4)^2 = 25$, etc. El dominio de f es en este caso todo el espacio $\mathbb{R}^2$, pues

$f(x_1, x_2) = x_1^2 + x_2^2$ puede ser calculado para cualquier pareja de números reales $(x_1, x_2) \in \mathbb{R}^2$. El rango de f es el conjunto de números reales no negativos, pues $f(x_1, x_2) = x_1^2 + x_2^2 \geq 0$ $\forall (x_1, x_2) \in \mathbb{R}^2$. ■

Ejemplo 2. La función $f(x_1, x_2, x_3) = \sqrt{1 - x_1^2 - x_2^2 - x_3^2}$ asocia a la terna $(x_1, x_2, x_3) \in \mathbb{R}^3$ el número real $\sqrt{1 - x_1^2 - x_2^2 - x_3^2}$. De este modo se tiene que $f(1, 0, 0) = \sqrt{1 - (1)^2 - (0)^2 - (0)^2} = 0 =$ imagen bajo f del vector $(1, 0, 0)$, $f(1/2, 1/2, 1/3) = \sqrt{1 - (1/2)^2 - (1/2)^2 - (1/3)^2} = \sqrt{7/18}$, etc. Obsérvese que entonces en este caso el dominio de f no es todo el espacio $\mathbb{R}^3$, por ejemplo $f(1, 1, 1) = \sqrt{1 - 1 - 1 - 1} = \sqrt{-2}$ que no existe (*como número real*, y estamos trabajando con funciones cuyas imágenes son números reales). Se tiene, de hecho, que

$$\begin{aligned} \text{dominio de } f &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 | 1 - x_1^2 - x_2^2 - x_3^2 \geq 0\} \\ &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1^2 + x_2^2 + x_3^2 \leq 1\} \end{aligned}$$ ■

NOTA. Cuando se trabaja con funciones de 2, 3 ó 4 variables, es usual, en lugar de usar subíndices con la letra x para denotar a las variables, usar las letras x y sus vecinas. Así, la función del ejemplo 1 anterior la escribimos como $f(x, y) = x^2 + y^2$, y la del ejemplo 2 como $f(x, y, z) = \sqrt{1 - x^2 - y^2 - z^2}$. Debemos advertir, sin embargo, que una letra como la z juega frecuentemente un doble papel: en funciones de dos variables es común usarla para denotar los valores de la función, escribiendo $z = f(x, y)$, (aquí, la z es "la función", como la del ejemplo 1: $z = x^2 + y^2$) en tanto que en funciones de 3 o más variables puede aparecer denotando a una de las variables de la función, por ejemplo, en $u = f(x, y, z)$ (aquí, la z es una variable independiente, como en la función del ejemplo 2: $u = \sqrt{1 - x^2 - y^2 - z^2}$). De cualquier modo, esto no debe provocar confusión.

Ejemplo 3. La función $f(x, y, z, w) = \frac{x^2 - y^2 - z^2 - w^2}{x^2 + y^2 + z^2 + w^2}$ es una función de las cuatro variables x, y, z, w que tiene por dominio a todo el espacio $\mathbb{R}^4$ excepto el $(0, 0, 0, 0)$. ■

En el caso concreto de funciones $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definidas en algún conjunto U del plano $\mathbb{R}^2$, conviene a veces tener una representación geométrica del dominio U de f. Este conjunto puede quedar determinado por las restricciones que f imponga para obtener $z = f(x, y) \in \mathbb{R}$ (el dominio natural de f), o bien, estar restringido por condiciones que arbitrariamente (por razones que hagan que así convenga) se impongan sobre él. En todo caso, estaremos interesados en "ver" el "pedazo" del plano $\mathbb{R}^2$ donde f está definida.

Ejemplo 4. La función $f(x, y) = \sqrt{1 - x^2 - y^2}$ está definida (su dominio natural es) el conjunto

$$U = \{(x, y) | 1 - x^2 - y^2 \geq 0\} = \{(x, y) | x^2 + y^2 \leq 1\}$$

el cual corresponde al interior del círculo unitario $x^2 + y^2 = 1$, incluyendo su frontera, como se muestra en la figura 2. Obsérvese que el rango de esta función es el conjunto $\{z \in \mathbb{R} | 0 \leq z \leq 1\}$ (¿por qué?). ■

Ejemplo 5. La función $f(x, y) = \ln(xy)$ está definida en aquellos puntos (x, y) en que $xy > 0$, restricción que impone la función logaritmo, pues ésta existe para valores positivos de su argumento.

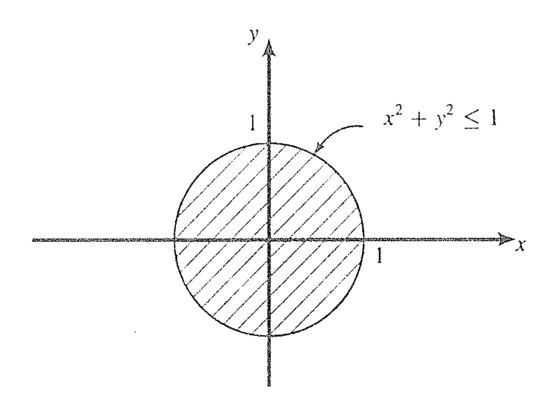


Figura 2. Dominio de la función $f(x, y) = \sqrt{1 - x^2 - y^2}$

Este conjunto U de puntos del plano con el producto de sus coordenadas positivo, queda descrito en forma más explícita como la unión de U_1 y U_2 donde

$$U_1 = \{(x, y)|x > 0, y > 0\}, \qquad U_2 = \{(x, y)|x < 0, y < 0\}$$

que corresponden a los puntos del primer y tercer cuadrante del plano xy, respectivamente.

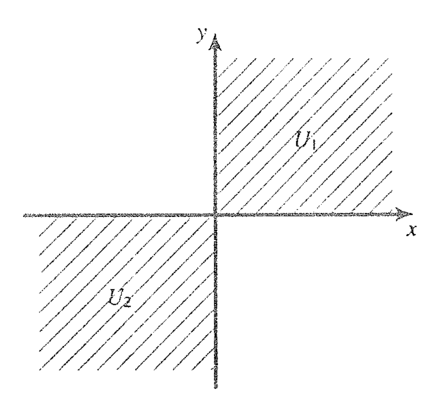


Figura 3. El dominio de la función $f(x, y) = \ln(xy)$.

Observe que el rango de esta función es todo el conjunto $\mathbb{R}$. De hecho, para puntos (x, y) en que $0 < xy < 1$, se tiene $f(x, y) < 0$, en tanto que si $xy > 1$, se tendrá $f(x, y) \geq 0$. Para los puntos de la hipérbola $xy = 1$, se tiene que $f(x, y) = 0$. ■

Ejemplo 6. La función $f(x, y) = \operatorname{arcsen} \frac{x}{x+y}$ tiene por dominio al conjunto de puntos (x, y) en el plano de tal modo que $-1 \leq \frac{x}{x+y} \leq 1$, restricción impuesta por la función arcsen, la cual, como sabemos, tiene por dominio al (rango de la función sen que es el) intervalo $[-1, 1]$. Es interesante describir explícitamente este conjunto, para poder tener una imagen geométrica de la región del plano que representa. Se trata entonces de "resolver" (dejar a la variable y en términos de x) la desigualdad

$$\left|\frac{x}{x+y}\right| \leq 1$$

Observando que de antemano el origen $x = y = 0$ está excluido del dominio de la función, se tiene que

$$\left|\frac{x}{x+y}\right| \leq 1 \Leftrightarrow |x| \leq |x+y| \Leftrightarrow x^2 \leq (x+y)^2$$

$$\Leftrightarrow y(y+2x) \geq 0 \Leftrightarrow \begin{cases} y \geq 0, & y \geq -2x \\ & \text{ó} \\ y \leq 0, & y \leq -2x \end{cases}$$

Así pues, el dominio U queda descrito por la unión de los subconjuntos U_1 y U_2 tales que

$$U_1 = \{(x, y)|y \geq 0, y \geq -2x\}, \qquad U_2 = \{(x, y)|y \leq 0, y \leq -2x\}$$

exceptuando el origen. Es decir, $U = (U_1 \cup U_2) - \{(0, 0)\}$. Geométricamente U_1 contiene los puntos que están *por encima* de la recta $y = -2x$ que tienen a su vez su ordenada no negativa, mientras que U_2 contiene los puntos que están *por debajo* de la recta $y = -2x$ que tienen su ordenada no positiva. El conjunto U se ve como

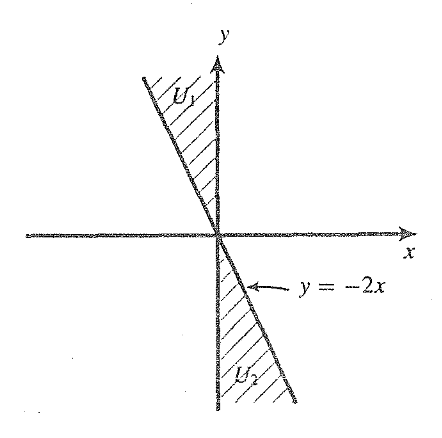


Figura 4. El dominio de la función $f(x, y) = \text{arcsen}\,\frac{x}{x+y}$.

El rango de la función f es el intervalo $[-\pi/2, \pi/2]$ (¿por qué?). ■

Las funciones reales de 3 variables, $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$, tienen por dominio a "pedazos" de $\mathbb{R}^3$, los cuales todavía pueden ser visualizados, como en el siguiente ejemplo.

Ejemplo 7. La función $f(x, y, z) = \ln(1 - x^2 - y^2 + z)$, tiene por dominio al conjunto

$$U = \{(x, y, z)|1 - x^2 - y^2 + z > 0\} = \{(x, y, z)|z > x^2 + y^2 - 1\}$$

La ecuación $z = x^2 + y^2 - 1$ representa un paraboloide que abre hacia arriba y tiene su vértice en el punto $(0, 0, -1)$ (detalles a este respecto se considerarán a partir de la siguiente sección). La región U de $\mathbb{R}^3$ queda entonces descrita como los puntos $(x, y, z) \in \mathbb{R}^3$ que están "dentro" del paraboloide mencionado (figura 5).

La igualdad entre dos funciones de varias variables se ve como sigue: las funciones $f_1: U_1 \subseteq \mathbb{R}^n \to \mathbb{R}$, $f_2: U_2 \subseteq \mathbb{R}^n \to \mathbb{R}$, se dicen ser *iguales*, lo cual se escribe como $f_1 = f_2$, si $U_1 = U_2$ y $f_1(x) = f_2(x)\ \forall x \in U_1$. ■

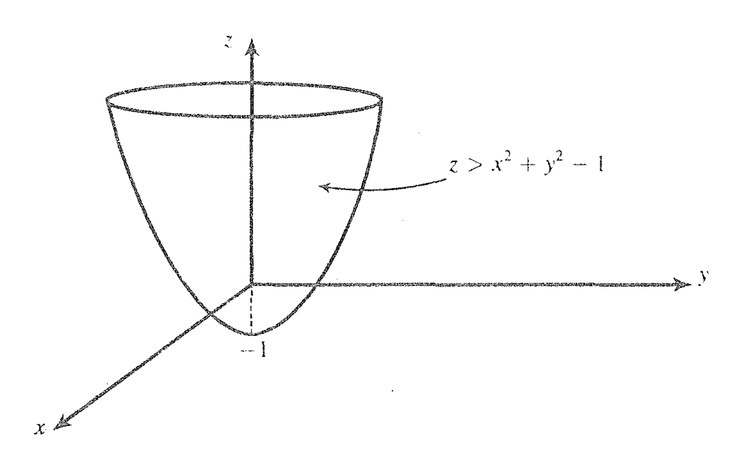


Figura 5. El dominio de la función $f(x, y, z) = \ln(1 - x^2 - y^2 + z)$.

Ejemplo 8. Las funciones $f(x, y) = |x + y|$ y $g(x, y) = |x| + |y|$ no son iguales, pues aunque sus dominios coinciden (es todo el espacio $\mathbb{R}^2$), sus imágenes no (en general). De hecho, sabemos que $f(x, y) = g(x, y)$ si y sólo si una de las coordenadas de (x, y) es un múltiplo no negativo de la otra (ver ejercicio 7 de la sección 3, capítulo 1). Por otra parte, las funciones $\phi(x, y) = \left|\frac{x}{y}\right|$ y $\psi(x, y) = \frac{|x|}{|y|}$ sí son iguales, pues sus dominios coinciden ($\mathbb{R}^2$ excepto el los puntos (x, y) con $y = 0$), así como sus imágenes. ■

Las operaciones entre funciones de varias variables se definen de manera análoga a como se hace en el caso de funciones de una sola variable: si tenemos las funciones $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ y $g: V \subseteq \mathbb{R}^n \to \mathbb{R}$, se define

a. la *suma de* f y g como la función $f + g: U \cap V \subseteq \mathbb{R}^n \to \mathbb{R}$, tal que

$$(f + g)(x) = f(x) + g(x)$$

b. el *producto de* f y g como la función $fg: U \cap V \subseteq \mathbb{R}^n \to \mathbb{R}$, tal que

$$(fg)(x) = f(x)g(x)$$

c. el *cociente de* f *entre* g como la función $\frac{f}{g}: W \subseteq \mathbb{R}^n \to \mathbb{R}$, tal que

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

donde $W = U \cap V - \{x \in V | g(x) = 0\}$.

Ejemplo 9. Sean $f(x, y) = \sqrt{1 - x^2 - y^2}$, $g(x, y) = \ln(xy)$, las funciones de los ejemplos 4 y 5 respectivamente. Sabemos que

$$\text{dominio de } f = \{(x, y) | x^2 + y^2 \leq 1\}$$
$$\text{dominio de } g = \{(x, y) | x > 0, y > 0\} \bigcup \{(x, y) | x < 0, y < 0\}$$

La intersección de estos dominios está entonces constituida por los puntos interiores al círculo unitario $x^2 + y^2 = 1$ (incluyendo su frontera) que están el el primer y tercer cuadrantes. Este es el dominio de la función suma de f y g,

$$(f + g)(x, y) = f(x, y) + g(x, y) = \sqrt{1 - x^2 - y^2} + \ln(xy)$$

y de la función producto de f y g,

$$(fg)(x, y) = f(x, y)g(x, y) = \sqrt{1 - x^2 - y^2}\ln(xy)$$

En el caso de la función cociente de f entre g, debemos eliminar de la intersección de los dominios de f y g los puntos (x, y) del dominio de g que hagan cero a esta función (es decir, debemos eliminar los ceros del denominador de la función cociente). Estos puntos son los de la hipérbola $xy = 1$, los cuales están fuera de la intersección de los dominios de f y g. Así pues, el dominio de la función cociente de f entre g

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} = \frac{\sqrt{1 - x^2 - y^2}}{\ln(xy)}$$

es el mismo que el de la función suma y producto.

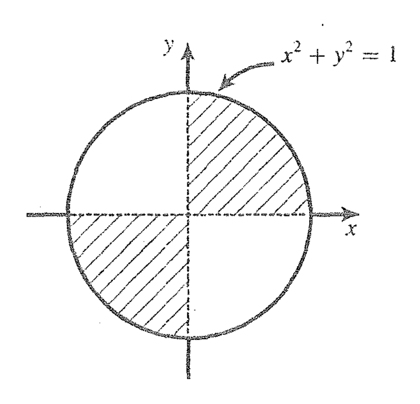


Figura 6. Los dominios de las funciones $f + g$, fg, f/g del ejemplo 9 ■

Para terminar esta sección, presentamos algunos tipos de funciones de varias variables a los que se les conoce con un nombre especial.

(*) A una función $f\colon \mathbb{R}^n \to \mathbb{R}$ del tipo $f(x) = c$, se le llama *función constante.*

(*) A una función $f\colon \mathbb{R}^n \to \mathbb{R}$ del tipo

$$f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$$

en la que $b \in \mathbb{R}$ y $a_1, a_2, \ldots, a_n$ son números reales dados no todos nulos, se le llama *función lineal.*

(*) A una función $f\colon \mathbb{R}^2 \to \mathbb{R}$ del tipo

$$f(x, y) = \sum_{i,j} a_{ij}x^i y^j$$

donde a_{ij} son números reales dados, se le llama *función polinomial* (en x y y). Por ejemplo, son funciones polinomiales $f(x, y) = x^2 + y^2 + 3$, $g(x, y) = 2 + 3x - 4y + 2x^2 - 5y^2 + xy$, etc. En general, una función polinomial en las n variables $x_1, x_2, \ldots, x_n$ es una función del tipo

$$f(x_1, x_2, \ldots, x_n) = \sum_{i_1, i_2, \ldots, i_n} a_{i_1 i_2 \ldots i_n} (x_1)^{i_1} (x_2)^{i_2} \ldots (x_n)^{i_n}$$

en la que $a_{i_1 i_2 \ldots i_n}$ son números reales dados. Por ejemplo, las funciones $f(x, y, z) = 3xy - 8xz$, $g(x, y, z, u) = 23 - 3z + 8x^2yz^3 - 2xyu^4 + u^5$, son funciones polinomiales.

Ejercicios (Capítulo 2, Sección 1)

1. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = x^2 + y^2$. Halle $f(1, 0)$, $f(0, 1)$, $f(1, 1)$. ¿Cuáles puntos $(x, y) \in \mathbb{R}^2$ son tales que $f(x, y) = 0$? ¿A dónde manda la función f los puntos (x, y) del círculo unitario $x^2 + y^2 = 1$?

2. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = x + y$. Halle $f(2, 3)$, $f(x, 1)$, $f(1, y)$, $f(x^{-1}, y^{-1})$. ¿Cuáles puntos $(x, y) \in \mathbb{R}^2$ son tales que $f(x, y) = k$? ¿A dónde manda f los puntos de la recta $y = -x$?

3. La función $f: \mathbb{R}^2 \to \mathbb{R}$ es tal que $f(x + y, x - y) = x^2 + y^2$. Determine $f(2, 5)$, $f(x, 3)$, $f(5, y)$, $f(x, y)$. ¿A dónde manda f los puntos de la recta $y = x$?, ¿a dónde manda f los puntos de la recta $y = -x$?

4. La función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es tal que $f(x - y, \frac{y}{x}) = y^2 - x^2$. Determine $f(x, y)$. ¿Cuál es el dominio U de esta función?

5. Sea $f: \mathbb{R}^3 \to \mathbb{R}$ la función $f(x, y, z) = e^{-(x^2+y^2+z^2)}$. Calcule $f(0, 0, 0)$ y $f(\pm 1, \pm 1, \pm 1)$. ¿A dónde manda f los puntos de la esfera unitaria $x^2 + y^2 + z^2 = 1$? ¿Qué pasa con los valores de f cuando $\|(x, y, z)\| \to \infty$?

6. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = \operatorname{sgn}(x - y^2)$, donde la función sgn (signo) está dada por

$$\operatorname{sgn}(\alpha) = \begin{cases} +1 & \text{si } \alpha > 0 \\ 0 & \text{si } \alpha = 0 \\ -1 & \text{si } \alpha < 0 \end{cases}$$

Describa el conjunto de puntos $(x, y) \in \mathbb{R}^2$ tales que: **a** $f(x, y) > 0$; **b** $f(x, y) = 0$; **c** $f(x, y) < 0$.

7. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = \ln(\operatorname{sgn}(1 + x^2 + y^2))$. ¿Cuál es el dominio de f?, ¿cuál es su rango?

8. Repita el ejercicio anterior con la función $f(x, y) = \operatorname{sgn}(\ln(1 + x + y))$.

En cada uno de los ejercicio 9 a 28, describa el dominio natural de la función $z = f(x, y)$ dada y haga un esquema en el que se represente este dominio en el plano xy.

9. $f(x, y) = \sqrt{x + y}$

10. $f(x, y) = \dfrac{1}{\sqrt{x + y}}$

11. $f(x, y) = \sqrt{x} + \sqrt{y}$

12. $f(x, y) = \dfrac{1}{\sqrt{x}} + \dfrac{1}{\sqrt{y}}$

13. $f(x, y) = \sqrt{xy}$

14. $f(x, y) = \sqrt{x}\sqrt{y}$

15. $f(x, y) = \sqrt{x + \sqrt{y}}$

16. $f(x, y) = \dfrac{1}{\sqrt{1 - x^2 - y^2}}$

17. $f(x, y) = \ln(1 + 2x^2 + 4y^2)$

18. $f(x, y) = \dfrac{1}{\sqrt{\ln(1 + 2x^2 + 4y^2)}}$

19. $f(x, y) = \sqrt{\ln(1 + x + y)}$

20. $f(x, y) = \ln(y \ln(1 + x + y))$

21. $f(x, y) = \sqrt{1 - x^2 - 5y^4}$

22. $f(x, y) = \arccos(x + y)$

23. $f(x, y) = \operatorname{arcsen}(x^2 + y)$

24. $f(x, y) = \arctan \dfrac{1 + x^2}{1 + y^2}$

25. $f(x, y) = \sqrt{\operatorname{senh}(2x + y)}$

26. $f(x, y) = \sqrt{\cosh(2x + y)}$

27. $f(x, y) = \sqrt{\operatorname{sen}[\pi(x^2 + y^2)]}$

28. $f(x, y) = \sqrt{y \cos x}$

En los ejercicios 29–33 haga explícito el dominio de la función indicada

29. $f(x, y, z) = \sqrt{1 + x^2 + y^2 + z^2}$

30. $f(x, y, z) = \sqrt{x + y + z} + \sqrt{x} + \sqrt{y} + \sqrt{z}$

31. $f(x, y, z) = \ln(x^4 \ln^2(y^2 \ln z))$

32. $f(x_1, x_2, x_3, x_4) = x_1 + 2x_2^2 + 3x_3^3 + 4x_4^4$

33. $f(x_1, x_2, x_3, x_4) = \dfrac{1}{\sqrt{x_1 + x_2 + x_3 + x_4}} + \sqrt{x_1 + x_2 + x_3 + x_4}$

En los ejercicios 34–40 diga si las funciones dadas son iguales

34. $f(x, y) = \sqrt{x/y}$, $g(x, y) = \sqrt{x}/\sqrt{y}$

35. $f(x, y) = |xy|$, $g(x, y) = |x||y|$

36. $f(x, y) = \ln(x + y)^2$, $g(x, y) = 2\ln(x + y)$

37. $f(x, y) = \ln(xy)$, $g(x, y) = \ln x + \ln y$

38. $f(x, y) = \ln(x^2/y^4)$, $g(x, y) = \ln x^2 - \ln y^4$

39. $f(x, y) = |\operatorname{sgn}(x + y)|$, $g(x, y) = \operatorname{sgn}|x + y|$

40. $f(x, y) = \operatorname{sgn}^2(1 - x - 4y)$, $g(x, y) = \operatorname{sgn}^{28}(1 - x - 4y)$

En los ejercicio 41–44 determine las funciones $f + g$, fg, f/g, así como sus dominios

41. $f(x, y) = x^2 + y^2$, $g(x, y) = 1$

42. $f(x, y) = x + y$, $g(x, y) = x - y$

43. $f(x, y) = \sqrt{1 + x + y}$, $g(x, y) = x - y$

44. $f(x, y, z) = \operatorname{sen}(x + y + z)$, $g(x, y, z) = 2\cos(x + y + z)$

45. ¿Cuál es el rango de la función constante $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = c$?

46. ¿Es una función constante un tipo especial de una función lineal?

47. ¿Es una función lineal un tipo especial de una función polinomial?

48. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto U de $\mathbb{R}^n$ y sea λ un número real. Defina la función $\lambda f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ como $(\lambda f)(\mathbf{x}) = \lambda f(\mathbf{x})$. Esta función es llamada "producto de la función f por el escalar λ". Obsérvese que éste es un caso particular de la operación producto de dos funciones, cuando una de ellas es una función constante. Si $g, h: U \subseteq \mathbb{R}^n \to \mathbb{R}$ son otras dos funciones definidas en U, demuestre que:

a. $f + g = g + f$

b. $f + (g + h) = (f + g) + h$

c. Existe una función $0: \mathbb{R}^n \to \mathbb{R}$ tal que $f + 0 = f$

d. Dada f, existe una función $-f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ tal que $f + (-f) = 0$ (la función cero del inciso anterior)

e. $\lambda(f + g) = \lambda f + \lambda g$, $\lambda \in \mathbb{R}$

f. $(\lambda + \mu)f = \lambda f + \mu f$, $\lambda, \mu \in \mathbb{R}$

g. $\lambda(\mu f) = \mu(\lambda f) = (\lambda\mu)f$, $\lambda, \mu \in \mathbb{R}$

h. $1f = f$

Así queda demostrado que el conjunto de todas las funciones reales definidas en $U \subseteq \mathbb{R}^n$ es un espacio vectorial (real) con las operaciones de suma (definida en esta sección) y producto por escalares (definida en este ejercicio).

(*) **49.** ¿Existen funciones inyectivas $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definidas en un abierto U de $\mathbb{R}^2$?

2.2 Geometría de las funciones de varias variables

Para funciones reales de una variable real $f: I \subseteq \mathbb{R} \to \mathbb{R}$, tenemos una representación geométrica de ellas, a saber, su gráfica (figura 1).

En el caso de una función de n variables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, definimos (análogamente al caso de una función de una variable) la *gráfica* de f como el conjunto

$$\{(x_1, x_2, \ldots, x_n, y) \in \mathbb{R}^{n+1} | (x_1, x_2, \ldots, x_n) \in U,\ y = f(x_1, x_2, \ldots, x_n)\}$$

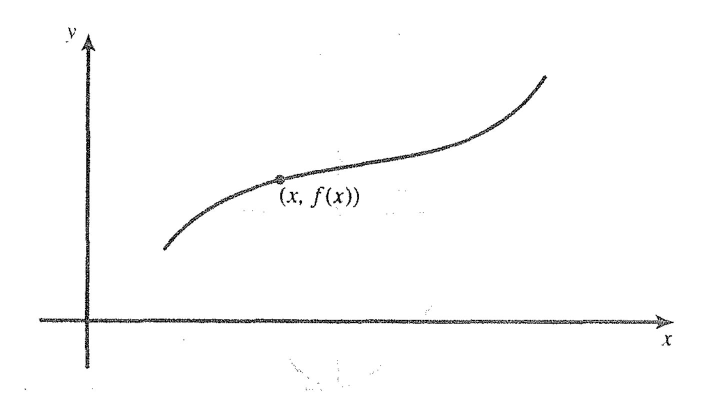


Figura 1. La gráfica de la función $y = f(x)$

Obsérvese que éste es un subconjunto del espacio $\mathbb{R}^{n+1}$. Así, si $n = 1$, tenemos la gráfica de la función $f: U \subseteq \mathbb{R} \to \mathbb{R}$ definida como el subconjunto del espacio $\mathbb{R}^2$

$$\{(x, y) \in \mathbb{R}^2 | x \in U, \ y = f(x)\}$$

tal como se dijo al principio de la sección.

Si $n = 2$, la gráfica de la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es el subconjunto del espacio $\mathbb{R}^3$

$$\{(x, y, z) \mid (x, y) \in U, z = f(x, y)\}$$

Podemos visualizar este subconjunto del espacio $\mathbb{R}^3$ como una "superficie en el espacio", poniendo los puntos de $U \subseteq \mathbb{R}^2$ en el plano $z = 0$ y los valores de $z = f(x, y)$ correspondientes en el eje z.

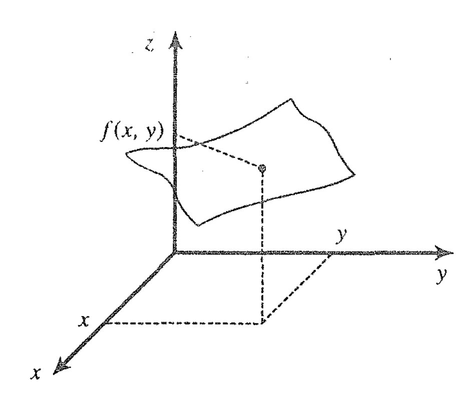


Figura 2. Gráfica de una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$

En otras palabras, la gráfica de la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, es la gráfica de la superficie $z = f(x, y)$.

Para $n \geq 3$, la gráfica de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ ya no puede ser visualizada, pues se encuentra en el espacio $(n + 1)$-dimensional ($\geq$ 4-dimensional). Es por este hecho que todos los conceptos importantes (por ejemplo, los relacionados a la diferenciabilidad) sobre funciones de varias variables, se ejemplifican de preferencia con funciones de 2 variables, pues en este caso tenemos representaciones geométricas concretas de las gráficas de ellas, donde es posible "ver" el concepto estudiado.

Ejemplo 1. La gráfica de la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2$ es un paraboloide cuyo aspecto es

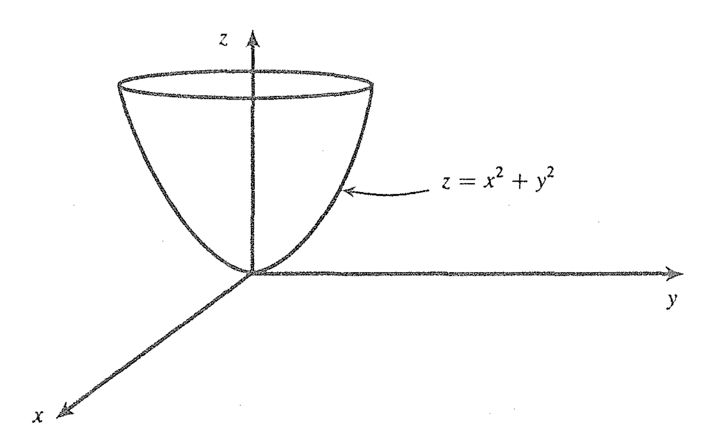


Figura 3. Gráfica de $f(x, y) = x^2 + y^2$. ■

Ejemplo 2. La gráfica de la función constante $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = k$, es un plano paralelo al plano xy, separado de éste una distancia de k unidades (hacia arriba si $k > 0$ o hacia abajo si $k < 0$).

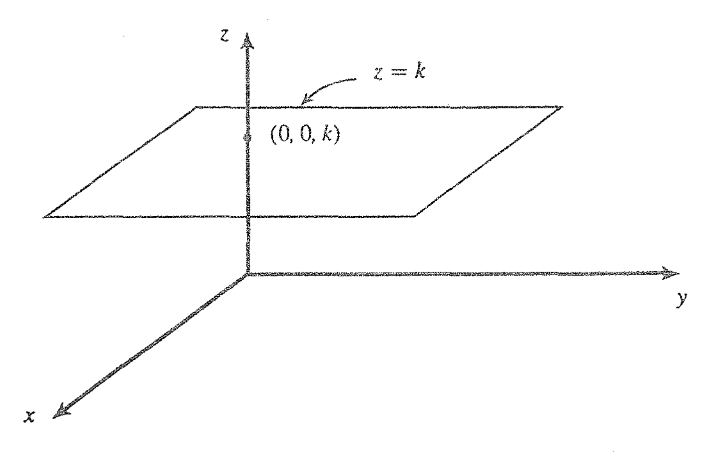


Figura 4. Gráfica de la función constante $f(x, y) = k$. ■

Ejemplo 3. La gráfica de la función lineal $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = ax + by + c$, es un plano que pasa por el punto $(0, 0, c)$ y un vector normal al plano es $(a, b, -1)$, (ó $(-a, -b, 1)$ como el mostrado en la figura 5). El ejemplo 2 es un caso particular de este ejemplo con $a = b = 0$.

Nótese la analogía con el caso de la función lineal $f: \mathbb{R} \to \mathbb{R}$ $f(x) = mx + b$. Esta es una recta en el plano $\mathbb{R}^2$ que pasa por $(0, b)$ y su vector normal es $(m, -1)$ (figura 6).
Una generalización de la función lineal de este ejemplo es la función lineal $f: \mathbb{R}^n \to \mathbb{R}$ del tipo

$$f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$$

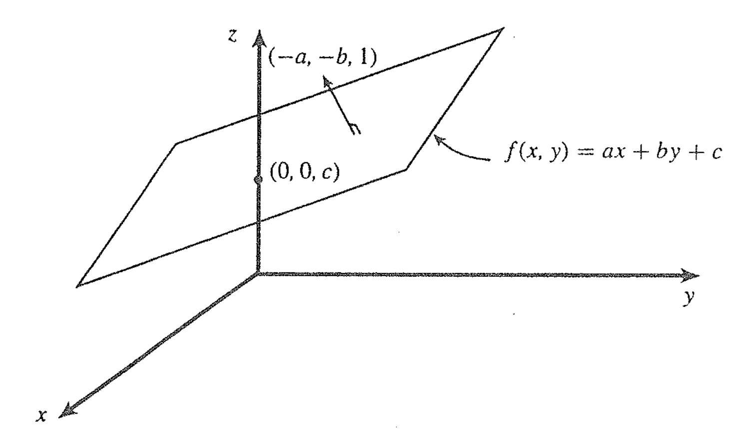


Figura 5. Gráfica de la función lineal $f(x, y) = ax + by + c$.

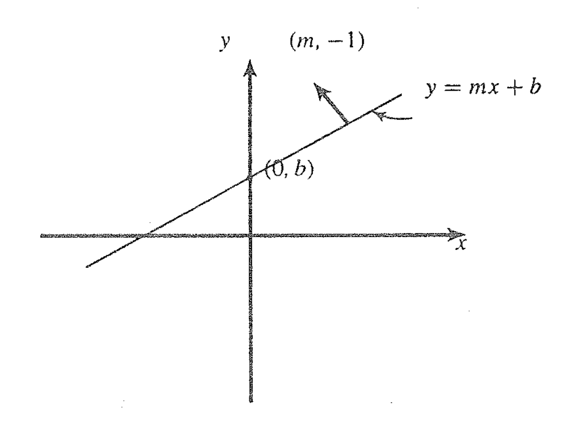


Figura 6. La función lineal de una variable $f(x) = mx + b$.

cuya gráfica es entonces el subconjunto de $\mathbb{R}^{n+1}$ dado por

$$\{(x_1, x_2, \ldots, x_n, y) | y = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b\}$$

Esta gráfica, aunque no podamos verla, decimos que es un "hiperplano" en $\mathbb{R}^{n+1}$ que pasa por $(0, 0, \ldots, 0, b)$ y cuyo vector normal es $(a_1, a_2, \ldots, a_n, -1)$. ■

Ejemplo 4. Recordemos que la ecuación $x^2 + y^2 = r^2$ no representa (globalmente) función alguna $f: I \subseteq \mathbb{R} \to \mathbb{R}$, pues al tratar de despejar "y" nos queda $y = \pm\sqrt{r^2 - x^2}$ de donde vemos que para cada valor de $x \in I = [-r, r]$, tenemos un par de valores asociados para "y", $\pm\sqrt{r^2 - x^2}$ con lo cual se viola la definición de función. Sabemos sin embargo que cada una de las expresiones $f_1(x) = \sqrt{r^2 - x^2}$, $f_2(x) = -\sqrt{r^2 - x^2}$ sí representan funciones del tipo $f_{1,2}: [-r, r] \to \mathbb{R}$, cuyas gráficas son los semicírculos superior e inferior, respectivamente, del círculo $x^2 + y^2 = r^2$, con centro en el origen y radio r. Del mismo modo la expresión $x^2 + y^2 + z^2 = r^2$ no representa (globalmente) función alguna $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, pues al tratar de escribir $z = f(x, y)$ nos quedaría, igual que en el caso anterior, $z = \pm\sqrt{r^2 - x^2 - y^2}$. Empero, las expresiones $f_1(x, y) = \sqrt{r^2 - x^2 - y^2}$,

$f_2(x, y) = -\sqrt{r^2 - x^2 - y^2}$ sí representan dos funciones del tipo $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ en donde $U = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq r^2\}$, cuyas gráficas son la semiesfera superior e inferior, respectivamente, de la esfera $x^2 + y^2 + z^2 = r^2$, con centro en el origen y radio r.

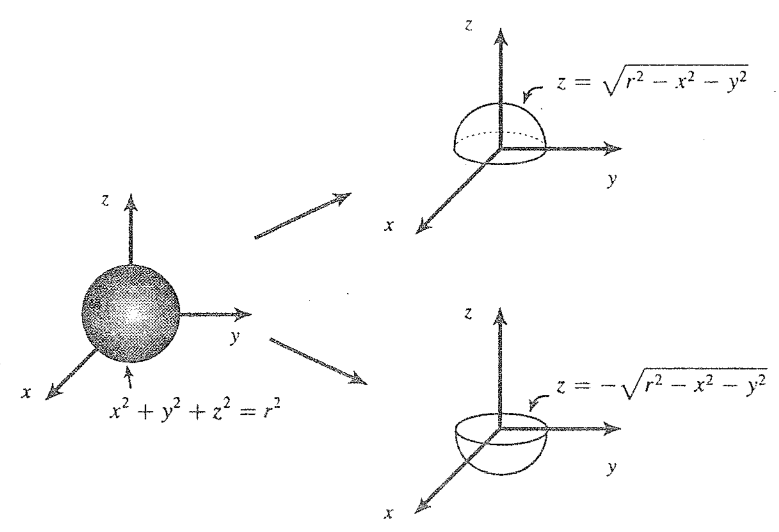


Figura 7. La esfera $x^2 + y^2 + z^2 = r^2$ y las funciones f_1 y f_2 que de ella se derivan. ■

Yendo más lejos, decimos que la "esfera en $\mathbb{R}^{n+1}$", con centro en el origen y radio r (cuando $r = 1$ es denotada por S^n) dada por

$$x_1^2 + x_2^2 + \cdots + x_n^2 + z^2 = r^2$$

define dos funciones del tipo $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, donde

$$U = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n | x_1^2 + x_2^2 + \cdots + x_n^2 \leq r^2\}$$

a saber

$$f_1(x_1, x_2, \ldots, x_n) = \sqrt{r^2 - x_1^2 - x_2^2 - \cdots - x_n^2}$$

$$f_2(x_1, x_2, \ldots, x_n) = -\sqrt{r^2 - x_1^2 - x_2^2 - \cdots - x_n^2}$$

cuyas gráficas son las "semiesfera superior e inferior" de la esfera dada respectivamente. ■

Es común identificar a una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ con su gráfica. Así, en algunas ocasiones nos referimos (en base a los ejemplos anteriores) a:

(*) "el paraboloide $f(x, y) = x^2 + y^2$".

(*) "el plano $f(x, y) = ax + by + c$".

(*) "la semiesfera superior $f(x, y) = \sqrt{r^2 - x^2 - y^2}$", etc.

En muchas ocasiones, dibujar (o simplemente bosquejar) la gráfica de una función de dos variables resulta una tarea altamente complicada. Por eso hay algunos procedimientos alternativos que nos

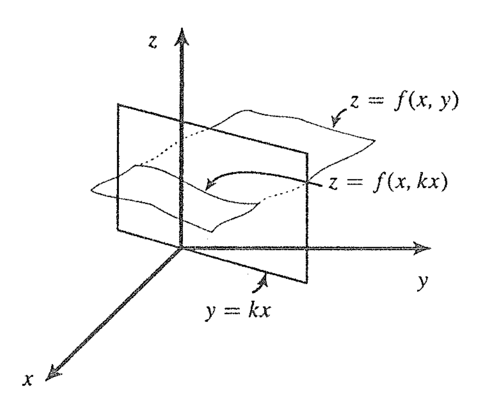


Figura 8. Intersección con planos $y = kx$.

permiten tener una descripción geométrica parcial de la gráfica de la función. Uno de ellos consiste en cortar la superficie $z = f(x, y)$ con planos del tipo $y = kx$ y estudiar el tipo de curvas así resultantes (las curvas $z = f(x, kx)$). En particular, se estudian las curvas resultantes de cortar la superficie con los planos xz ($y = 0$) y yz ($x = 0$).

Ejemplo 5. Consideremos la función $z = f(x, y) = x^2 + y^2$. Cuando $y = 0$, nos queda la curva $z = x^2$ (en el plano zx), y cuando $x = 0$ nos queda la curva $z = y^2$ (en el plano zy). Ambas son parábolas con vértice en el origen que abren hacia arriba. Más aún, si $y = kx$, nos queda $z = (1 + k^2)x^2$ que también es una parábola de las mismas características que las anteriores.

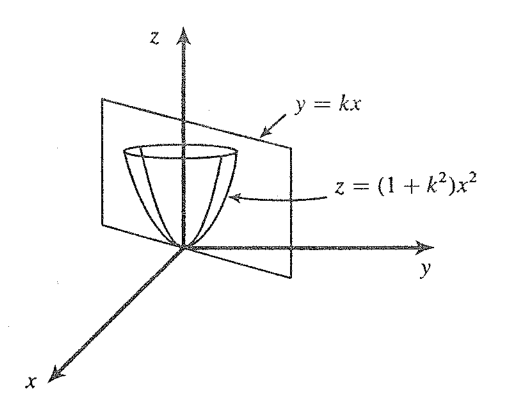


Figura 9. Intersección del paraboloide $f(x, y) = x^2 + y^2$ con los planos $y = kx$.

■

Ejemplo 6. La gráfica de la función $z = 1 - x - y$ es un plano. Cuando $y = 0$, nos queda la recta $z = 1 - x$. En general, cortando con $y = kx$, nos queda la recta $z = 1 - (1 + k)x$ (figura 10).

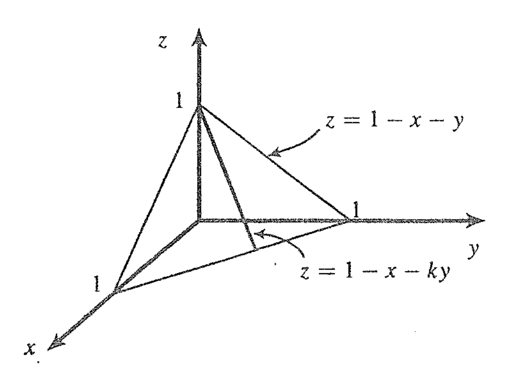


Figura 10. Intersección del plano $z = 1 - x - y$ con $y = kx$. ■

Otro concepto importante que también se usa como ayuda para tener representaciones geométricas parciales de gráficas de funciones de dos variables, es el de "nivel constante" de una función: dada la función $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, y el número $c \in$ rango de f, se define el *nivel c de* la *función f* como el conjunto de puntos de U que f manda a c. Es decir

$$\{x \in U \mid f(x) = c\} = f^{-1}(c)$$

Esquemáticamente

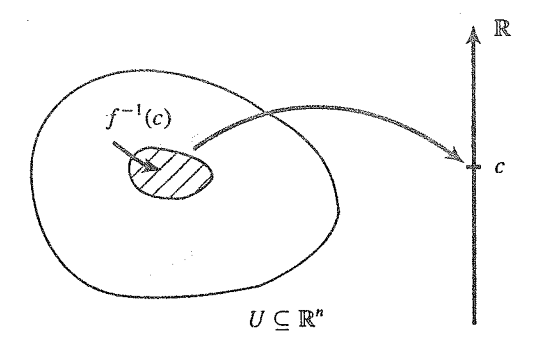


Figura 11. Nivel c de f.

Nótese que éste es un subconjunto del espacio $\mathbb{R}^n$, de modo que cuando $n = 1$, es un subconjunto de números reales; cuando $n = 2$, es un subconjunto del espacio $\mathbb{R}^2$, el cual es en muchas ocasiones una curva llamada *curva de nivel* (de la función f correspondiente al nivel $z = c$). Cuando $n = 3$, el nivel c de la función $f: U \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ es un subconjunto del espacio $\mathbb{R}^3$, el cual es también, en muchas ocasiones, una superficie llamada *superficie de nivel.*

Ejemplo 7. Considere la función $f: \mathbb{R} \to \mathbb{R}$, $f(x) = \operatorname{sen} x$. El nivel 1 de esta función está constituido por

$$f^{-1}(1) = \{x \in \mathbb{R} | \operatorname{sen} x = 1\} = \{x = \frac{\pi}{2} + 2k\pi, k \in \mathbb{Z}\}$$

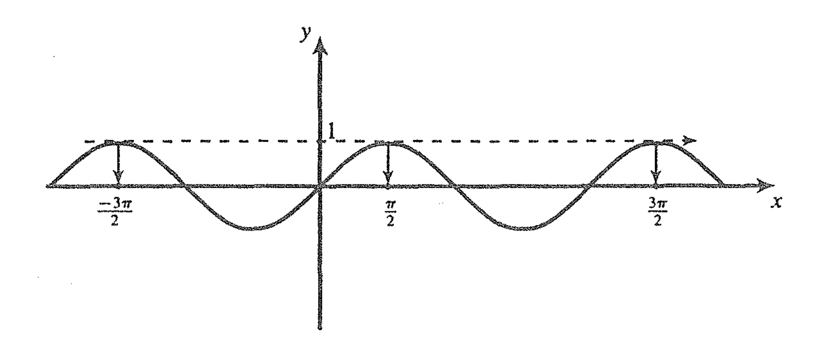


Figura 12. El nivel 1 de la función $f(x) = \operatorname{sen} x$.

En este caso el nivel 1 está constituido por una infinidad de puntos aislados. Para la función $f: \mathbb{R} \to \mathbb{R}$, $f(x) = \operatorname{sgn} x$ tenemos que el nivel 1 es

$$f^{-1}(1) = \{x \in \mathbb{R} | \operatorname{sgn} x = 1\} = \{x \in \mathbb{R} | x > 0\}$$

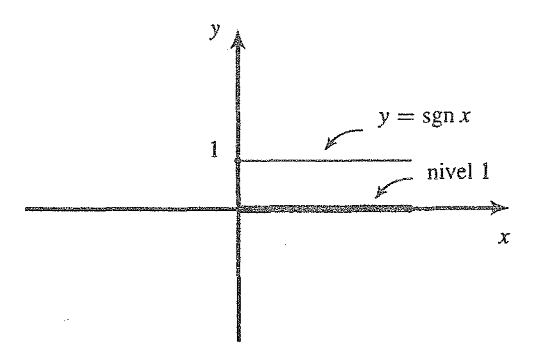


Figura 13. El nivel 1 de la función $f(x) = \operatorname{sgn} x$.

En este caso el nivel 1 está formado por todas las x de la parte positiva de la recta real. Para la función $f: \mathbb{R} \to \mathbb{R}$, $f(x) = 2x + 3$, el nivel 1 es

$$f^{-1}(x) = \{x \in \mathbb{R} | 2x + 3 = 1\} = \{x = -1\}$$

En este caso el nivel 1 está formado por un único punto (figura 14).

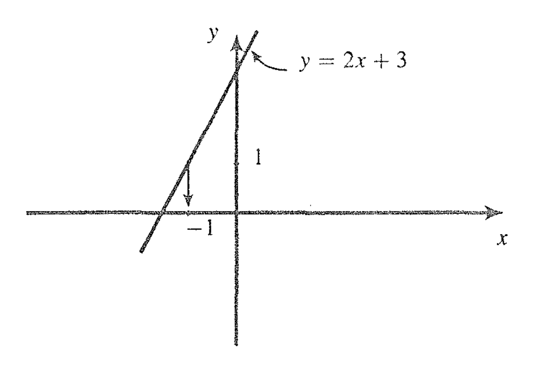


Figura 14. El nivel 1 de la función $f(x) = 2x + 3$. ■

Nótese también que el nivel c de la superficie $z = f(x, y)$ se puede interpretar geométricamente como la intersección de dicha superficie con el plano $z = c$ (al resolver simultáneamente las ecuaciones $z = f(x, y)$ y $z = c$ se obtiene $f(x, y) = c$), la cual es, como dijimos, (en algunos casos) una curva en el plano.[1] Esta curva (de nivel) es $f(x, y) = c$, la cual, de manera natural, se identifica como una curva *del plano xy*. Obsérvese, sin embargo, que bajo la perspectiva de la intersección de la superficie con el plano $z = c$, esta curva se debe encontrar en el *plano* $z = c$ (pues al resolver simultáneamente el sistema $z = f(x, y)$, $z = c$, se obtienen las ternas $(x, y, z) \in \mathbb{R}^3$ tales que $f(x, y) = c$, y $z = c$. De este modo la curva $f(x, y) = c$ que dibujamos en el plano xy es la proyección en este plano de la intersección de la superficie $z = f(x, y)$ con el plano $z = c$.

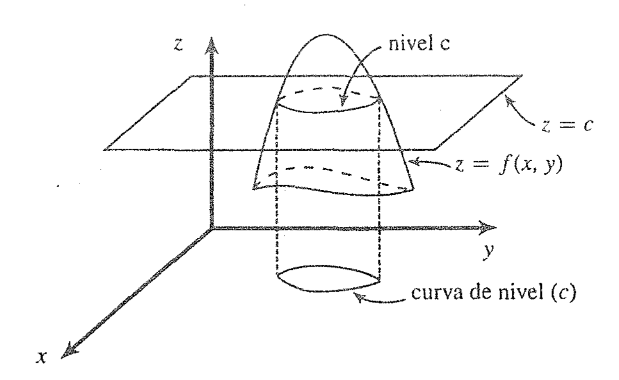


Figura 15. La intersección de $z = f(x, y)$ con $z = c$.

Ejemplo 8. Consideremos nuevamente la función del ejemplo 1, $f(x, y) = x^2 + y^2$. Nótese que $f(x, y) \geq 0 \ \forall (x, y) \in \mathbb{R}^2$, de modo que el rango de f está formado por los reales no negativos. Sea

[1] En este momento no podemos decir en qué casos la intersección de la gráfica de la función $z = f(x, y)$ con el plano $z = c$ produce una curva en el plano, pues, entre otras cosas, no sabemos aún lo que en matemáticas se entiende por "curva". Este es un importante concepto con el que se trabajará ampliamente en el capítulo 5. Aun identificando "curva" con "gráfica de una función real de variable real", la cuestión de cuándo la intersección mencionada produce una curva es un asunto no trivial que se discutirá en el capítulo 3.

pues $c \geq 0$. El nivel c de la función f es el conjunto

$$f^{-1}(c) = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = c\}$$

Si $c = 0$, éste se reduce al punto $(0, 0)$. Si $c > 0$, $f^{-1}(c)$ es un círculo con centro en el origen y radio $\sqrt{c}$. Estos círculos constituyen las curvas de nivel de la función f.

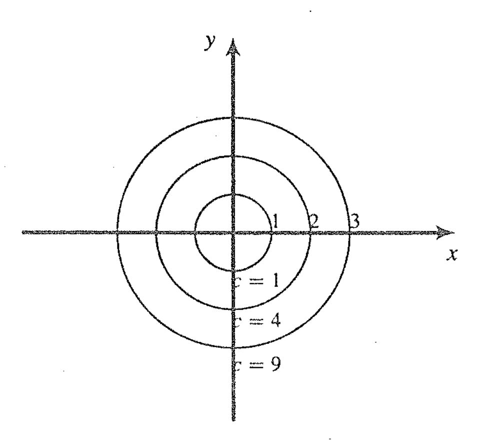


Figura 16. Curvas de nivel de la función $f(x, y) = x^2 + y^2$. ■

Ejemplo 9. Sea $f(x, y) = \sqrt{x^2 + y^2}$. Las curvas de nivel de esta función son los círculos $x^2 + y^2 = c^2$, que tienen el mismo aspecto de las curvas de nivel de la función del ejercicio anterior. Nótese que, sin embargo, la curva de nivel 2 de la función del ejercicio anterior no corresponde a la curva del nivel 2 de esta función.

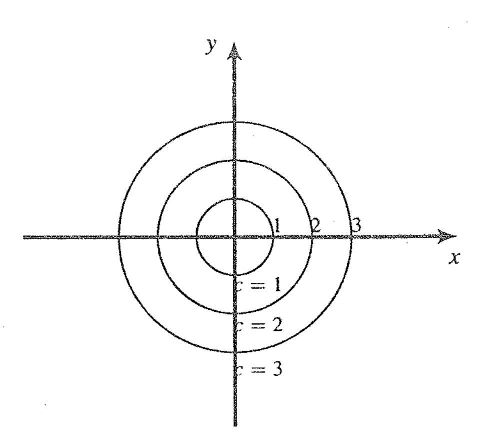


Figura 17. Curvas de nivel de la función $f(x, y) = \sqrt{x^2 + y^2}$.

Continuando con este ejemplo, si usamos las ideas expuestas en los ejemplos 5 y 6 anteriores, vemos que con $y = 0$ obtenemos la curva $z = \sqrt{x^2} = |x|$, y con $x = 0$ la curva $z = \sqrt{y^2} = |y|$. Más aún, intersectando con el plano $y = kx$ obtenemos la curva $z = \sqrt{1 + k^2}|x|$. Es fácil ver, con esto, que la superficie $z = \sqrt{x^2 + y^2}$ es un cono con vértice en el origen y que abre hacia arriba (figura 18).

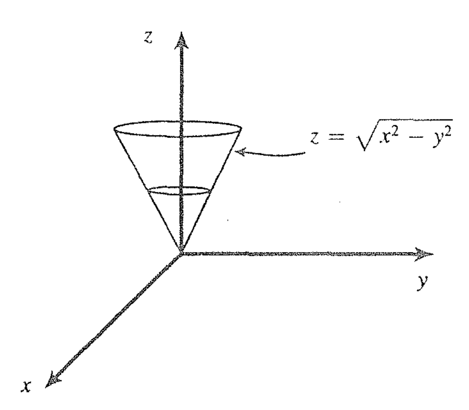


Figura 18. Intersecciones del cono $z = \sqrt{x^2 + y^2}$ con $z = c$ y $y = kx$. ■

Ejemplo 10. La función $f: \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$, tiene por nivel c $(c > 0)$ a

$$f^{-1}(c) = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R} | x_1^2 + x_2^2 + \cdots + x_n^2 = c\}$$

Obsérvese que si $n = 2$, este es el ejemplo 8 anterior. Con $n = 3$, la función $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$, tiene por superficie de nivel a la esfera $x_1^2 + x_2^2 + x_3^2 = c$. ■

Ejemplo 11. Consideremos, como último ejemplo, la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^2$. Esta es una función interesante que nos servirá posteriormente para ejemplificar algunos conceptos que aparecerán en nuestro estudio de la diferenciabilidad de funciones de varias variables. Téngase presente en principio que el rango de f es todo $\mathbb{R}$. Si $c > 0$, las curvas de nivel de esta función están constituidas por hiperbolas del tipo $x^2 - y^2 = c$, en tanto que si $c < 0$, las hiperbolas son del tipo $y^2 - x^2 = -c$. Para $c = 0$, tenemos $x^2 - y^2 = (x - y)(x + y) = 0$, que son dos rectas que se cortan en el origen, a saber $y = x$, $y = -x$.

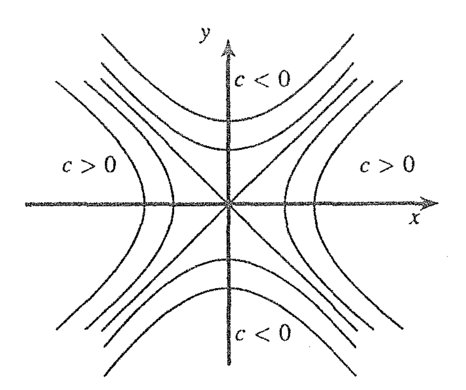


Figura 19. Curvas de nivel de la función $f(x, y) = x^2 - y^2$.

Véase también que la intersección de la superficie $z = x^2 - y^2$ con $y = 0$ es la parábola $z = x^2$, que tiene vértice en el origen y *abre hacia arriba*, en tanto que si intersectamos tal superficie con $x = 0$,

obtenemos la parábola $z = -y^2$, cuyo vértice está en el origen y *abre hacia abajo*. Esta superficie presenta pues un comportamiento interesante alrededor del origen. Su forma es la de una "silla de montar". Con este nombre nos referiremos a ella en adelante.

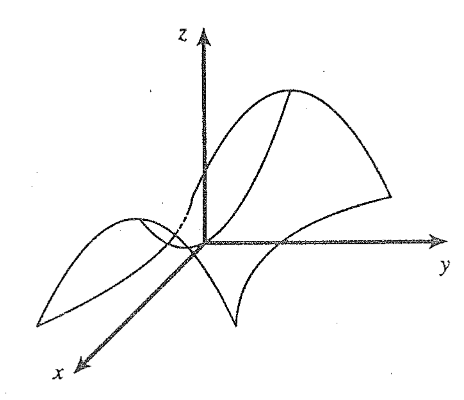


Figura 20. Gráfica de la función $f(x, y) = x^2 - y^2$. ■

Ejercicios (Capítulo 2, Sección 2)

1. Considere la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $z = f(x, y)$.
   a. Sea $g(x, y) = f(x, y) + k$, en donde k es un número real dado. ¿Dónde se define la función g? ¿Cómo es la gráfica de la función g respecto de la gráfica de la función f? Ejemplifique con la función $g(x, y) = x^2 + y^2 + 3$.
   b. Sea $g(x, y) = f(x - x_0, y - y_0)$, donde (x_0, y_0) es un punto de $\mathbb{R}^2$ dado. ¿Dónde se define la función g? ¿Cómo es la gráfica de la función g respecto de la gráfica de la función f? Ejemplifique con la función $g(x, y) = (x - 2)^2 + (y + 3)^2$.
   c. Sea $g(x, y) = f(-x, -y)$. ¿Dónde se define la función g? ¿Cómo es la gráfica de la función g respecto de la gráfica de la función f? Ejemplifique con la función $g(x, y) = (-x - 2)^2 + (-y + 3)^2$ (usando la gráfica del inciso anterior).
   d. Sea $g(x, y) = -f(x, y)$. ¿Dónde se define la función g? ¿Cómo es la gráfica de la función g respecto de la gráfica de la función f? Ejemplifique con la función $g(x, y) = -x^2 - y^2$.

2. Para cada una de las funciones lineales $f: \mathbb{R}^2 \to \mathbb{R}$ dadas a continuación, dé un vector normal al plano que representa la función, así como un punto por el que éste pasa.
   a. $f(x, y) = 3x - 2y + 3$ **d.** $f(x, y) = y$
   b. $f(x, y) = -x + y - 1$ **e.** $f(x, y) = x + 7$
   c. $f(x, y) = x$ **f.** $f(x, y) = 2$

3. Dé un ejemplo de una función $f: \mathbb{R} \to \mathbb{R}$
   a. cuyo nivel 1 sea $\{x = 1\}$
   b. cuyo nivel 1 sea $\{x = 1, x = 3\}$

c. cuyo nivel 1 sea $\{1 \leq x \leq 3\}$

d. cuyo nivel 1 sea $\{x | x \in \mathbb{Z}\}$

e. cuyo nivel 1 sea $\mathbb{R}$.

4. ¿Verdadero o falso? La función $f: I \subseteq \mathbb{R} \to \mathbb{R}$, definida en $I \subseteq \mathbb{R}$ es biyectiva si y sólo si para cada $c \in \mathbb{R}$ el nivel c de f consta de un solo punto.

5. En el texto se dijo que el nivel constante de una función de dos variables $z = f(x, y)$ era "en muchas ocasiones" una curva. Digamos que si $c \in$ rango de f, la curva sería $f(x, y) = c$. Dé un ejemplo en el que $f(x, y) = c$ *no* es una curva en el plano. Ahora bien, dé un ejemplo de una función $z = f(x, y)$ y un valor $c \in$ rango de f, de modo que:

a. $f(x, y) = c$ sea un punto;

b. $f(x, y) = c$ sea el primer cuadrante del plano xy, sin incluir a los ejes.

c. $f(x, y) = c$ sea todo el plano $\mathbb{R}^2$.

6. Describa las curvas de nivel de una función lineal $f(x, y) = ax + by + c$.

7. A la gráfica de la función $f(x, y) = ax^2 + by^2$ se le llama "paraboloide elíptico". ¿Cómo son las curvas de nivel de esta función?

8. Se define la función "mínimo de dos números", mín: $\mathbb{R}^2 \to \mathbb{R}$, como

$$\text{mín}(x, y) = \begin{cases} x & \text{si } x \leq y \\ y & \text{si } y < x \end{cases}.$$

Obtenga mín$(1, -1)$, mín$(3, \pi)$, mín$(3, e)$. Demuestre que la gráfica de esta función es simétrica respecto del plano $y = x$. Describa las curvas de nivel de esta función.

9. Se define la función "máximo de dos números", máx: $\mathbb{R}^2 \to \mathbb{R}$, como

$$\text{máx}(x, y) = \begin{cases} x & \text{si } x \geq y \\ y & \text{si } y > x \end{cases}.$$

Obtenga máx$(2^{1/2}, 2^{3/2})$, máx$(2^{-1/2}, 2^{-3/2})$. Demuestre que la gráfica de esta función es simétrica respecto del plano $y = x$. Describa las curvas de nivel de esta función. ¿A qué es igual la función $z = \text{mín}(\text{máx}(x, y), \text{mín}(x, y))$?

10. a. Describa las curvas de nivel de la función $z = \text{mín}(|x|, |y|)$.

b. Describa las curvas de nivel de la función $z = \text{máx}(|x|, |y|)$.

c. Describa las curvas de nivel de la función $z = \text{mín}(x^2, y)$.

11. Dé un ejemplo de una función $f: \mathbb{R}^2 \to \mathbb{R}$

a. cuyo nivel 1 sea la curva $y = \operatorname{sen} x$.

b. cuyo nivel -7 sea la curva $y = \sqrt{x^6 + \ln^8 x}$.

c. cuyo nivel 126 sea la curva $y^4 x + x^3 y - 5 = 0$.

d. cuyo nivel 0 sea el conjunto de puntos del interior del círculo unitario $x^2 + y^2 = 1$ (sin incluir la frontera).

12. Sea $\phi: I \subseteq \mathbb{R} \to \mathbb{R}$ una función real de una variable real, con rango $J \subseteq \mathbb{R}$. Dé una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ cuyo nivel c sea la gráfica de la función ϕ. ¿Dónde se define la función f?

13. Considere la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, y sean α y β dos números de su rango ($\alpha \neq \beta$). Sean $C_1 = f^{-1}(\alpha)$, $C_2 = f^{-1}(\beta)$ los niveles constantes correspondientes a α y β. Viendo a C_1 y C_2 como subconjuntos del plano xy, demuestre que éstos tienen intersección vacía (es decir, dos curvas de nivel de una función de dos variables no se pueden intersectar).

14. Sea $f(x, y) = (y - x^2 + 1)(y + x^2 - 1)$. Demuestre que el nivel cero de esta función está formado por las curvas $y = x^2 - 1$, $y = 1 - x^2$, *las cuales* se *cortan* en $(1, 0)$ y en $(-1, 0)$. ¿Contradice esto el resultado del ejercicio anterior? Explique.

15. Discuta la siguiente afirmación: el nivel cero de una función $z = f(x, y)$ es "la manera" como dicha superficie "pasa por el plano xy".

En los ejercicios 16–24, describa las curvas de nivel de las funciones indicadas. Haga una gráfica mostrando algunas de estas curvas.

16. $f(x, y) = |x| - y$

17. $f(x, y) = x - |y|$

18. $f(x, y) = |x - y|$

19. $f(x, y) = \sqrt{xy}$

20. $f(x, y) = \dfrac{x}{y}$

21. $f(x, y) = \dfrac{2x}{x^2 + y^2}$

22. $f(x, y) = \dfrac{2y}{x^2 + y^2}$

23. $f(x, y) = (\operatorname{sgn} x)y$

24. $f(x, y) = \operatorname{arcsen}(x + y)$

25. En el texto se dijo que el nivel constante de una función de tres variables $u = f(x, y, z)$ era "en muchas ocasiones" una superficie. Digamos que si $c \in$ rango de f, la superficie sería $f(x, y, z) = c$. Dé un ejemplo en el que $f(x, y, z) = c$ *no* es una superficie en el espacio. Más aún, dé un ejemplo de una función $u = f(x, y, z)$ y un valor $c \in$ rango de f, tal que:

a. $f(x, y, z) = c$ sea un punto;

b. $f(x, y, z) = c$ sea la línea $x = y = z$;

c. $f(x, y, z) = c$ sea el primer octante del espacio xyz, sin incluir los planos coordenados.

26. Describa las superficies de nivel de la función lineal $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = ax + by + cz + d$.

27. Dé una función $f: \mathbb{R}^3 \to \mathbb{R}$

a. cuyo nivel 1 sea la superficie $z = x^2 + y^2$.

b. cuyo nivel -7 sea la superficie $z = \ln^2(\operatorname{sen}^4(x + y^8) + 7)$.

c. cuyo nivel 126 sea la superficie $xz^3 + x^2y^5z^2 - 23yz + 128 = 0$.

d. cuyo nivel 0 sea el conjunto de puntos del interior de la esfera unitaria $x^2 + y^2 + z^2 = 1$ (sin incluir la frontera).

28. Sea $\phi: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el subconjunto U de $\mathbb{R}^2$, con rango $J \subseteq \mathbb{R}$. Sea c un punto de J. Dé una función $u = f(x, y, z)$ cuyo nivel c sea la superficie $z = \phi(x, y)$. ¿Cuál es el dominio de f?

29. Discuta la siguiente afirmación: el nivel cero de una función $u = f(x, y, z)$ es "la manera" como la hipersuperficie, gráfica de f, "atraviesa el espacio $\mathbb{R}^3$".

30. Considere la función $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$, y sean α y β dos números de su rango. Sean $S_1 = f^{-1}(\alpha)$, $S_2 = f^{-1}(\beta)$ los niveles constantes correspondientes a α y β ($\alpha \neq \beta$). Demuestre que S_1 y S_2 son disjuntos (es decir, dos superficies de nivel de una función de tres variables no se pueden intersectar). En general, si $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, y sean α y β dos números de su rango. Demuestre que $f^{-1}(\alpha) \cap f^{-1}(\beta) = \varnothing$

En los ejercicios 31–36, describa las superficies de nivel de las funciones $u = f(x, y, z)$ dadas.

31. $f(x, y, z) = x^2 + y^2 - z^2$

32. $f(x, y, z) = x^2 - y^2 + z^2$

33. $f(x, y, z) = x^2 - y^2 - z^2$

34. $f(x, y, z) = 2x^2 + 4y^2 + 10z^2$

35. $f(x, y, z) = \dfrac{x^2 + y^2}{z}$

36. $f(x, y, z) = x^2 + y^2 + z^2 - 4x - 6y + 15z$

37. (Superficies de revolución). Situémonos en el espacio $\mathbb{R}^3$. Piense en una curva en el semiplano superior del plano yz, digamos la gráfica de una función no negativa $z = f(y)$. Si ponemos a girar esta gráfica, teniendo como eje de giro el eje z, obtenemos una superficie S en $\mathbb{R}^3$, llamada *superficie de revolución.* Obsérvese que un punto $p = (0, y, z) \in \mathbb{R}^3$ de la gráfica de $z = f(y)$, girará en un círculo C alrededor del eje z. Este círculo deberá ser una curva de nivel de la superficie S. Es decir, *todos los puntos* (x, y) *del círculo* C (viéndolo proyectado en el plano xy), *deben tener* la *misma imagen* z. La distancia de un punto cualquiera (x, y) de C al origen es $\sqrt{x^2 + y^2}$. Es claro entonces que la función cuya gráfica es la superficie de revolución S es $z = f\left(\sqrt{x^2 + y^2}\right)$.

a. Considere la parábola $z = y^2$ (en el plano zy). Demuestre que la superficie de revolución que se obtiene al girar esta parábola alrededor del eje z es el paraboloide $z = x^2 + y^2$.

b. Considere la función valor absoluto $z = |y|$ (en el plano zy). Demuestre que la superficie de revolución que se obtiene al girar la gráfica de $z = |y|$ alrededor del eje z es el cono $z = \sqrt{x^2 + y^2}$.

c. Determine la superficie de revolución que se obtiene al girar el semicírculo superior $z = \sqrt{a^2 - y^2}$ alrededor del eje z.

d. Determine la superficie de revolución que se obtiene al girar la catenaria $z = \cosh y$ alrededor del eje z.

e. Demuestre que la superficie $z = e^{-(x^2+y^2)}$ es una superficie de revolución. ¿Qué aspecto tiene?

f. ¿Podría considerar un plano como superficie de revolución?

38. (Más superficies de revolución...)

a. Suponga que la gráfica de la función $y = f(z)$, dibujada sobre el plano zy, gira alrededor del eje y. Demuestre que se obtiene así una superficie de revolución cuya ecuación es $y = f\left(\sqrt{z^2 + x^2}\right)$.

b. Suponga que la gráfica de la función $x = f(z)$, dibujada sobre el plano zx, gira alrededor del eje x. Demuestre que se obtiene así una superficie de revolución cuya ecuación es $x = f\left(\sqrt{z^2 + y^2}\right)$.

c. ¿Las superficies descritas en los incisos **a** y **b** pueden ser gráficas de una función $z = \phi(x, y)$? Explique.

d. ¿Las superficies descritas en los incisos **a** y **b** pueden ser superficies de nivel de una función $u = \psi(x, y, z)$? Explique.

2.3 Límites y continuidad

Para poder abordar adecuadamente el estudio de la diferenciabilidad de funciones de varias variables es necesario tener algunos conceptos sobre limites y continuidad de estas funciones. No es objeto de esta obra un acercamiento riguroso y exhaustivo de esta parte de la teoría. Por una parte, establecer "rigurosamente" los resultados que, sobre limites y continuidad, aparecen en el estudio de las funciones de varias variables, requiere un trabajo previo con la topología del espacio $\mathbb{R}^n$, lo cual no debe interesar demasiado en un primer curso de cálculo (si el objetivo del curso fuera hacer "análisis", la historia sería otra). Por otra parte, el estudio de los límites y la continuidad poco ayuda, a posteriori, a entender las técnicas del cálculo en el espacio $\mathbb{R}^n$, siendo éstas uno de los objetivos importantes en un primer curso sobre esta materia.

La exposición que se presenta a continuación puede parecer poco ambiciosa para algunos, o quizá demasiado pretenciosa para otros. Para los últimos dejamos al profesor la decisión de tomar sólo alguna parte de este material, o sugerir la profundidad con que se debe estudiar. Para los primeros presentamos algunas demostraciones *opcionales* de varios resultados que aparecen en esta parte de la teoría, al mismo tiempo que se recomienda estudiar algunos de los excelentes libros de análisis matemático que existen actualmente en el mercado.

Al encontrarnos con el estudio de los límites de las funciones de varias variables, se ponen al descubierto las grandes dificultades de pasar del cálculo de una al de varias variables: para funciones de una variable sus dominios son "pedazos de la recta", muchas veces intervalos, cuyas descripciones son en general fáciles de entender. Para una función de n variables, su dominio es un "pedazo de $\mathbb{R}^n$", y... aquí empiezan los problemas. ¿Cómo son los subconjuntos de $\mathbb{R}^n$ "equivalentes" a los subconjuntos de $\mathbb{R}$? Esta es una pregunta cuya respuesta puede ser muy complicada , y pertenece a la parte de la matemática llamada "topología". Lo que a continuación haremos es estudiar solamente aquellos conceptos "topológicos" que nos hagan lo más eficiente posible el camino para comprender los conceptos de límite y continuidad para una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$.

Recordemos primeramente el concepto de límite para una función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ definida en el intervalo abierto I de $\mathbb{R}$. Sea $x_0 \in I$. Aceptamos que $f(x_0)$ puede no existir, es decir que f esté definida en $I - \{x_0\}$. Decir que el límite de $f(x)$, cuando x tiende a x_0, es L, significa que estando x cerca de x_0 se tiene $f(x)$ cerca de L. El concepto de "cercanía" en la recta lo podemos establecer rigurosamente con la idea de "vecindad": una vecindad de centro x_0 y radio r es el conjunto

$$\{x \in \mathbb{R} \mid |x - x_0| < r\} = \{x \in \mathbb{R} | x_0 - r < x < x_0 + r\}$$

es decir, en este conjunto están las x "vecinas a x_0 en menos que r". Con este concepto, decíamos que el límite de $f(x)$ es L, que se escribe $\lim_{x\to x_0} f(x) = L$, si siempre que se diera una vecindad con centro en L y radio $\epsilon > 0$ (en la recta de las imágenes de la función) uno podía conseguir una vecindad con centro en x_0 y radio $\delta > 0$ con la propiedad de que cualquier x que esté en esta vecindad (excepto posiblemente x_0) tuviese su imagen dentro de la vecindad dada. Es decir, $\lim_{x\to x_0} f(x) = L$ significa que dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$0 < |x - x_0| < \delta \Rightarrow |f(x) - L| < \epsilon$$

lo cual se ve geométricamente como (figura 1)

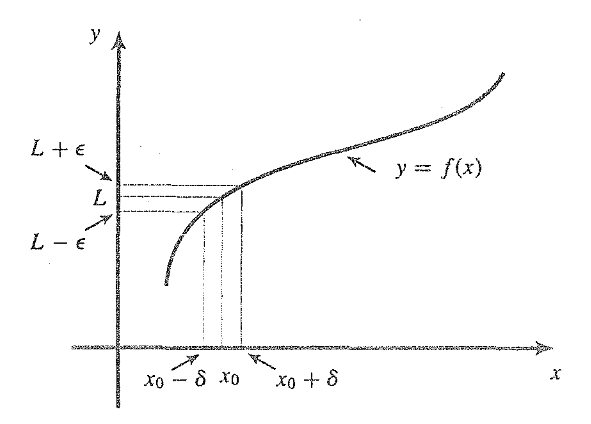


Figura 1. Interpretación geométrica de la expresión $\lim_{x\to x_0} f(x) = L$.

El concepto de límite para funciones de varias variables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ será, en esencia, el mismo. El problema que se nos presenta ahora es que cuando queramos hablar de "cercanía" para la x de la función (que vive en $\mathbb{R}^n$) las vecindades ya no nos servirán (para las imágenes de $f(x)$ seguirán sirviendo). Tendremos que comenzar entonces por introducir un concepto que generalice la idea de vecindad que teníamos en la recta, con el cual podamos establecer rigurosamente la frase "cuando $\mathbf{x} \in \mathbb{R}^n$ esté cerca de $\mathbf{x}_0 \in \mathbb{R}^n \ldots$". La herramienta básica que usaremos para nuestro propósito es la de "bola abierta", que se da en la siguiente definición.

Definición. (bola abierta) Sea $\mathbf{x}_0 \in \mathbb{R}^n$ y $r > 0$. La bola abierta de centro en $\mathbf{x}_0$ y radio r, denotada por $B(\mathbf{x}_0, r)$, es el conjunto de puntos de $\mathbb{R}^n$ que distan de $\mathbf{x}_0$ en menos que r. Es decir

$$B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| < r\}$$ ■

Recordemos que la norma (euclidiana) del vector $\mathbf{x} - \mathbf{x}_0$, que aparece en la definición anterior, es (por definición) la distancia entre $\mathbf{x}$ y $\mathbf{x}_0$. Es por eso que la desigualdad $\|\mathbf{x} - \mathbf{x}_0\| < r$ se cumple precisamente para aquellos puntos $x \in \mathbb{R}^n$ cuya distancia a $\mathbf{x}_0$ es menor que r.

Veamos cómo son las bolas abiertas en $\mathbb{R}$. En este caso

$$B(x_0, r) = \{x \in \mathbb{R} \mid |x - x_0| < r\} = \{x \in \mathbb{R} | x_0 - r < x < x_0 + r\}$$

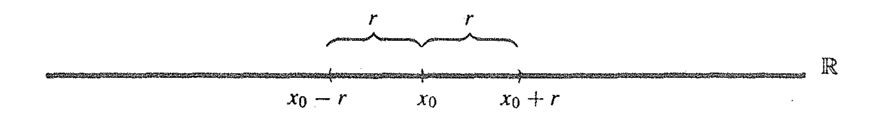


Figura 2. Una bola abierta en ℝ.

que geométricamente se ve como en la figura 2

Es decir, las bolas abiertas de ℝ son precisamente nuestras viejas conocidas "vecindades con centro en x_0 y radio r", que nos daban una descripción de las x "vecinas a x_0 en menos que r". Es precisamente esta misma idea la que recoge el concepto de "bola abierta en $\mathbb{R}^n$".

Para $n = 2$, escribiendo $\mathbf{x} = (x, y)$, $\mathbf{x}_0 = (x_0, y_0)$, tenemos

$$\begin{aligned} B(\mathbf{x}_0, r) &= \{(x, y) \in \mathbb{R}^2 \mid \|(x, y) - (x_0, y_0)\| < r\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid \|(x - x_0, y - y_0)\| < r\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid \sqrt{(x - x_0)^2 + (y - y_0)^2} < r\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 < r^2\} \end{aligned}$$

Es decir, la bola $B(\mathbf{x}_0, r)$, está constituida por los puntos interiores a la circunferencia $(x - x_0)^2 + (y - y_0)^2 = r^2$, o sea, los puntos que distan de (x_0, y_0) menos que r.

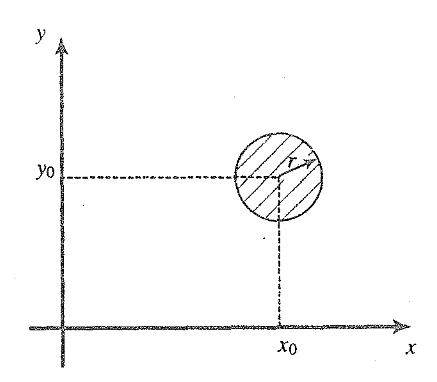


Figura 3. Bola abierta en $\mathbb{R}^2$.

Igualmente, para $n = 3$, si $\mathbf{x} = (x, y, z)$, $\mathbf{x}_0 = (x_0, y_0, z_0)$, tenemos

$$B(\mathbf{x}_0, r) = \{(x, y, z) \in \mathbb{R}^3 | (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 < r^2\}$$

que es un conjunto que describe a los puntos $\mathbf{x} \in \mathbb{R}^3$ vecinos de $\mathbf{x}_0$ en menos que r. Estos son los puntos interiores de una esfera con centro en $\mathbf{x}_0$ y radio r (figura 4).

Definamos ahora lo que se entiende por "conjunto abierto".

Definición. (conjunto abierto). Se dice que el conjunto $U \subseteq \mathbb{R}^n$ es un *conjunto abierto* de $\mathbb{R}^n$, si para cada $\mathbf{x}_0 \in U$ existe un $r > 0$ tal que $B(\mathbf{x}_0, r) \subset U$. ■

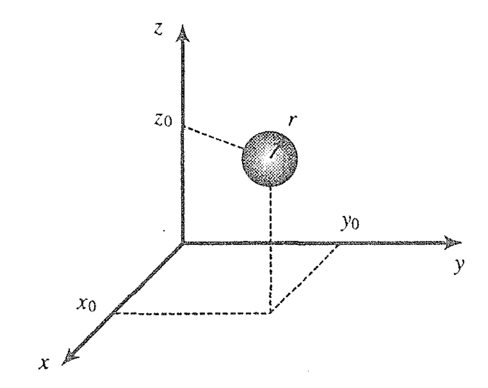


Figura 4. Bola abierta en $\mathbb{R}^3$.

Es decir, el conjunto $U \subseteq \mathbb{R}^n$ será abierto si cuando tomamos un punto $\mathbf{x}_0$ de él, este siempre tiene vecinos (en una bola abierta de centro en $\mathbf{x}_0$ y algún radio $r > 0$) que siguen viviendo dentro de U. Obsérvese que no importa qué tan grande o pequeña sea la bola donde viven los vecinos de $\mathbf{x}_0$. Lo que importa es que tales vecinos (que viven en la bola) existan. De hecho, es fácil imaginar que a medida que $\mathbf{x}_0$ esté más cerca de "la frontera de U" (si es que existe), los vecinos de la definición anterior vivirán en una bola cada vez más pequeña.

La idea intuitiva de un conjunto abierto es, pues, un conjunto "que no incluye su frontera"

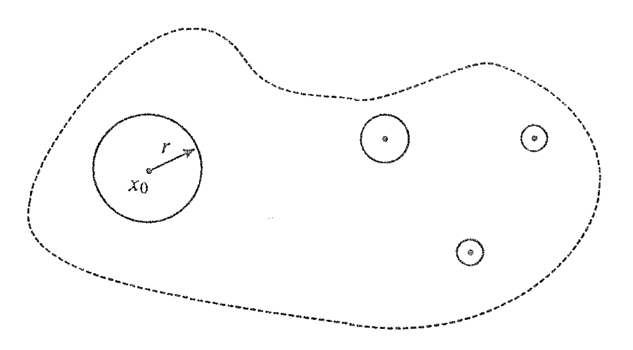


Figura 5. Un conjunto abierto.

Un intervalo abierto es un conjunto abierto de $\mathbb{R}$. Debe ser claro que todo el espacio $\mathbb{R}^n$ es un conjunto abierto. También el conjunto vacío $\varnothing$ es abierto (¿por qué?).

Ejemplo 1. Hay una infinidad de ejemplos de conjuntos abiertos. Daremos ahora algunos de ellos. Más que pensar en "cómo justificar" que el conjunto es abierto, es mejor desarrollar una cierta intuición que nos diga convincentemente "tal conjunto es abierto y tal otro no lo es". Una bola abierta es un conjunto abierto [demostración opcional: Considere la bola $B(\bar{\mathbf{x}}, s) \subset \mathbb{R}^n$ de centro en $\bar{\mathbf{x}} \in \mathbb{R}^n$ y radio $s > 0$. Para ver que este es un conjunto abierto tome cualquier $\mathbf{x}_0 \in B(\bar{\mathbf{x}}, s)$. Veremos que existe $r > 0$ tal que $B(\mathbf{x}_0, r) \subset B(\bar{\mathbf{x}}, s)$. De hecho, sea $r = s - \|\mathbf{x}_0 - \bar{\mathbf{x}}\|$. Observe que $r > 0$ pues

siendo $\mathbf{x}_0$ un elemento de $B(\bar{\mathbf{x}}, s)$, se tiene $\|\mathbf{x}_0 - \bar{\mathbf{x}}\| < s$. Tome cualquier $\mathbf{y} \in B(\mathbf{x}_0, r)$. Se tienen las siguientes implicaciones

$$\begin{aligned}
\mathbf{y} \in B(\mathbf{x}_0, r) &\Rightarrow \|\mathbf{y} - \mathbf{x}_0\| < r \Rightarrow \|\mathbf{y} - \mathbf{x}_0\| < s - \|\mathbf{x}_0 - \bar{\mathbf{x}}\| \\
&\Rightarrow \|\mathbf{y} - \mathbf{x}_0\| + \|\mathbf{x}_0 - \bar{\mathbf{x}}\| < s \\
&\Rightarrow \|\mathbf{y} - \bar{\mathbf{x}}\| \leq \|\mathbf{y} - \mathbf{x}_0\| + \|\mathbf{x}_0 - \bar{\mathbf{x}}\| < s \qquad \begin{pmatrix}\text{la primera desigualdad} \\ \text{es la desigualdad triangular}\end{pmatrix} \\
&\Rightarrow \mathbf{y} \in B(\bar{\mathbf{x}}, s)
\end{aligned}$$

Con esto se tiene entonces que $B(\mathbf{x}_0, r) \subset B(\bar{\mathbf{x}}, s)$. Como $\mathbf{x}_0 \in B(\bar{\mathbf{x}}, s)$ fue arbitrario concluimos que $B(\bar{\mathbf{x}}, s)$ es abierto]. Si A es un conjunto abierto de $\mathbb{R}^n$ y $C = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ es un conjunto finito de puntos de A, entonces $A - C = \{\mathbf{x} \in A, \mathbf{x} \notin C\}$ sigue siendo un conjunto abierto. El semiplano superior $\{(x, y) \in \mathbb{R}^2 | y > 0\}$ y el semiplano inferior $\{(x, y) \in \mathbb{R}^2 | y < 0\}$ son conjuntos abiertos de $\mathbb{R}^2$. El plano xy de $\mathbb{R}^3$, $\{(x, y, z) \in \mathbb{R}^3 | z = 0\}$ *no* es un conjunto abierto de $\mathbb{R}^3$. ■

Definición. (Frontera de un conjunto). Sea $U \subseteq \mathbb{R}^n$ un subconjunto de $\mathbb{R}^n$. Se dice que el punto $\mathbf{x}_0 \in \mathbb{R}^n$ es un punto frontera de U si toda bola abierta con centro en $\mathbf{x}_0$ y radio $r > 0$ contiene puntos dentro de U y fuera de U. La *frontera* de U es el conjunto de puntos frontera de U, y se denota por ∂U. ■

Téngase presente que siendo $\mathbf{x}_0 \in \mathbb{R}^n$ un punto frontera de U (el cual puede o no pertenecer a U) se debe tener *para todo* $r > 0$

$$B(\mathbf{x}_0, r) \cap U \neq \varnothing \quad \text{y} \quad B(\mathbf{x}_0, r) \cap U^c \neq \varnothing$$

Si U es un conjunto abierto, y si $\mathbf{x}_0$ es un punto frontera de U, entonces $\mathbf{x}_0$ *no puede pertenecer a* U (¿por qué?). En tal caso la segunda condición anterior se cumple automáticamente, pues $\{\mathbf{x}_0\} \subset B(\mathbf{x}_0, r) \cap U^c$, de tal modo pues que si U es abierto, el punto $\mathbf{x}_0$ ($\notin U$) es un punto frontera de U si y sólo si cualquier bola abierta con centro en $\mathbf{x}_0$ y radio $r > 0$ contiene puntos que pertenecen a U, i.e. si $B(\mathbf{x}_0, r) \cap U \neq \varnothing$ $\forall r > 0$.

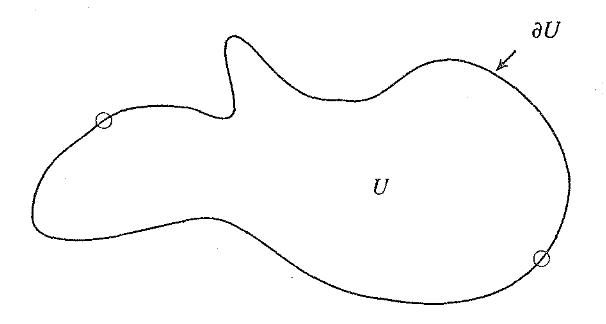


Figura 6. Frontera de un conjunto.

Ejemplo 2. Si I es un intervalo finito de $\mathbb{R}$, los extremos de él son puntos frontera de I. La frontera de la bola abierta $B(\mathbf{x}_0, r)$ de $\mathbb{R}^n$ (que ya vimos que es un conjunto abierto) es la esfera $\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n | x_1^2 + x_2^2 + \cdots + x_n^2 = r^2\}$. En $\mathbb{R}^2$ se ve como

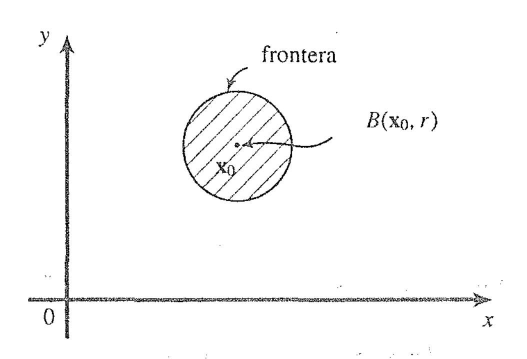


Figura 7. Una bola en $\mathbb{R}^2$ y su frontera.

Ahora definiremos el concepto de límite de una función de varias variables. Si $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función definida en el conjunto abierto U de $\mathbb{R}^n$ y $\mathbf{x}_0$ es un punto de U o bien un punto frontera de U, nos interesa estudiar el *comportamiento* de los valores de f para los valores de x que estén cerca de $\mathbf{x}_0$. *No* queremos estudiar lo que pasa con f en $\mathbf{x}_0$, pues estamos aceptando que f puede incluso no estar definida en $\mathbf{x}_0$ (si $\mathbf{x}_0$ es un punto frontera de U, siendo U abierto se tiene $\mathbf{x}_0 \notin U$, i.e. $\mathbf{x}_0$ no está dentro del dominio de la función f). Así, si $f(\mathbf{x}_0)$ existe, tal valor no nos va a interesar. Lo que queremos, pues, es estudiar qué pasa con los valores de $f(\mathbf{x})$ cuando $\mathbf{x}$ es una vecina "íntima" ($\mathbf{x}$ vive en una bola de centro $\mathbf{x}_0$ y radio $r > 0$, r pequeño) de $\mathbf{x}_0$. Estudiaremos, pues, expresiones del tipo

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(x) = L$$

donde $L \in \mathbb{R}$, $\mathbf{x}, \mathbf{x}_0 \in \mathbb{R}^n$, lo cual, de manera intuitiva (y quizás esta sea la manera más importante de entender este concepto), significa que siendo $\mathbf{x}$ una "vecina íntima" de $\mathbf{x}_0$, entonces $f(\mathbf{x})$ es una "vecina íntima" de L, o bien, que estando $\mathbf{x}$ suficientemente cerca de $\mathbf{x}_0$, se puede tener a $f(x)$ suficientemente cerca de L. Todavia más, de otro modo, que podemos tener a $f(\mathbf{x})$ tan cerca de L como queramos, con tal de que $\mathbf{x}$ esté lo suficientemente cerca de $\mathbf{x}_0$. Todas estas ideas intuitivas se cristalizan rigurosamente en la siguiente definición.

Definición. (límite) Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $\mathbf{x}_0$ un punto de U o bien un punto frontera de U. Se dice que el límite de f cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$ es L, lo cual se escribe como

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = L$$

si dado cualquier $\epsilon > 0$ (por pequeño que sea) existe $\delta > 0$ tal que

$$\mathbf{x} \in B(\mathbf{x}_0, \delta) \cap U (\mathbf{x} \neq \mathbf{x}_0) \Rightarrow f(\mathbf{x}) \in B(L, \epsilon)$$ ■

Esquemáticamente

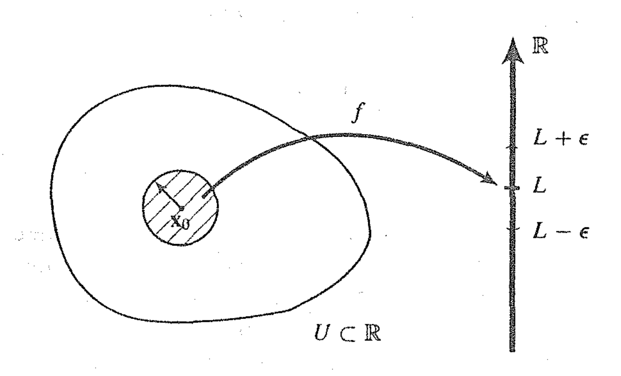


Figura 8. El concepto de límite.

Si $n = 2$ se tiene la siguiente interpretación geométrica

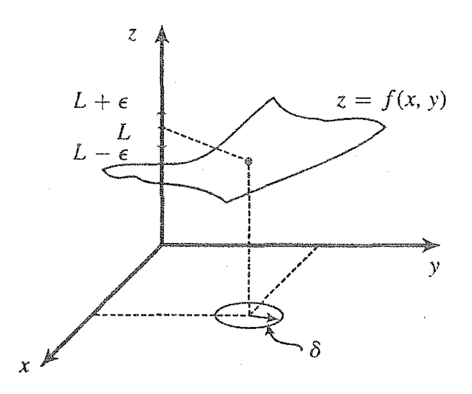


Figura 9. El concepto de límite para funciones de dos variables.

Ejemplo 3. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ la función constante $f(\mathbf{x}) = c$. Si $\mathbf{x}_0 \in \mathbb{R}^n$, entonces $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = c$. Este es un hecho elemental que se puede probar con la definición dada anteriormente. Sugerimos, sin embargo, que el lector piense con argumentos intuitivos, como por ejemplo: $f(\mathbf{x})$ es una función que manda a cualquier punto $\mathbf{x} \in \mathbb{R}^n$ siempre al mismo valor c. Preguntarse, bajo esta perspectiva, qué hace $f(\mathbf{x})$ cuando $\mathbf{x}$ es vecina íntima de $\mathbf{x}_0$ (algún $\mathbf{x}_0 \in \mathbb{R}^n$ dado) resulta trivial: $\mathbf{x}$ puede hacer "lo que le dé la gana" y $f(\mathbf{x})$ estará *siempre* en el valor de c. Así pues, si $\mathbf{x}$ es vecina íntima de $\mathbf{x}_0$, $f(\mathbf{x})$ será vecina íntima de c (aceptamos que la relación "ser vecinas íntimas" es reflexiva, i.e. c es vecina íntima de c para todo $c \in \mathbb{R}$). ∎

Ejemplo 4. Sea $f: \mathbb{R}^n \to \mathbb{R}$ la función

$$f(x_1, x_2, \ldots, x_n) = ax_i$$

donde a es un real dado no nulo e i es un índice fijo, $1 \leq i \leq n$ (ejemplos de estas funciones son: $f(x, y) = 2x$, $f(x, y, z) = 20z$, etc.). Sea $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ un punto dado de $\mathbb{R}^n$. Se tiene $\lim_{\mathbf{x}\to\bar{\mathbf{x}}} f(\mathbf{x}) = a\bar{x}_i$. Dejamos al lector que aplique la definición de límite para validar esta afirmación (tome $\delta = \epsilon/a$ y use que $|x_i - \bar{x}_i| \leq \|\mathbf{x} - \bar{\mathbf{x}}\|$). Así entonces se tiene

$$\lim_{(x,y)\to(3,8)} 2x = 2(3) = 6, \qquad \lim_{(x,y)\to(3,8)} 8y = 8(8) = 64,$$

$$\lim_{(x,y,z)\to(1,1,0)} 20z = 20(0) = 0, \text{ etc.}$$ ■

Ejemplo 5. Considere la función $f(x, y) = \frac{xy}{x^2+y^2}$. Esta función está definida en $\mathbb{R}^2 - \{(0, 0)\}$. Queremos estudiar el límite $\lim_{(x,y)\to(0,0)} f(x, y)$. Es decir, queremos ver cómo se comportan los valores $f(x, y)$ cuando (x, y) está cerca de (0,0). Una manera de tener "candidatos" al valor del límite (si éste existe) es hacer que (x, y) tienda a (0,0) por medio de un camino concreto dado por una curva $y = \phi(x)$ que pasa por el origen. En tal caso, si el límite $\lim_{(x,y)\to(0,0)} f(x, y)$ existe y vale L, y el límite $\lim_{x\to0} f(x, \phi(x))$ existe, éste debe valer L. La ventaja de este planteamiento es que el último límite es el de una función de una sola variable x, el cual, en principio, podemos calcular sin mucha dificultad. Observe cuidadosamente la estructura lógica de las afirmaciones hechas anteriormente: si el límite de $f(x, y)$ cuando (x, y) tiende a (0,0) existe y vale L, y si el límite $\lim_{x\to0} f(x, \phi(x))$ existe (donde $y = \phi(x)$ es una curva que pasa por el origen), éste debe valer L.

En el caso que nos ocupa, poniendo $y = 0$ (el eje x) tenemos

$$\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{x\to0} f(x, 0) = \lim_{x\to0} \frac{x(0)}{x^2 + 0^2} = 0$$

Esto *no* significa que $\lim_{(x,y)\to(0,0)} f(x, y) = 0$. Lo que significa es que si nos acercamos al origen por el eje x, los valores de la función $f(x, y)$ se acercan a 0 (más aún, $f(x, y)$ *vale* cero en los puntos del eje x).

Hagamos lo mismo por el camino $y = x$

$$\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{x\to0} f(x, x) = \lim_{x\to0} \frac{x(x)}{x^2 + x^2} = \frac{1}{2}$$

Es claro entonces que el límite $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ no existe: no puede ser 0 pues los puntos del tipo (x, x) con x muy pequeño son vecinos íntimos del origen y sin embargo $f(x, x)$ no es vecino íntimo del 0; por la misma razón no puede ser 1/2, pues los puntos $(x, 0)$, con x muy pequeño son vecinos íntimos del origen pero $f(x, 0)$ no es vecino íntimo de 1/2. Este ejemplo nos presenta un "procedimiento" para convencernos de que un límite $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ no existe (el cual no nos sirve para ver que el límite sí existe —¿por qué?). ■

Como resumen del ejemplo anterior diremos: una condición *necesaria* (no suficiente) para que el límite $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ exista y sea L, es que si los límites

$$\lim_{x\to x_0} f(x, \phi(x)) \qquad \lim_{x\to x_0} f(x, \psi(x))$$

existen (donde $y = \phi(x)$, $y = \psi(x)$ son curvas que pasan por (x_0, y_0)), deben valer L.

Ejemplo 6. Con las ideas del ejemplo 5 podemos ver que el límite $\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^4+y^2}$ no existe. El siguiente par de cálculos nos permite concluirlo

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^4+y^2} = \overset{y=0}{\longrightarrow} = \lim_{x\to 0} \frac{3x^2(0)}{x^4+(0)^2} = 0$$

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^4+y^2} = \overset{y=x^2}{\longrightarrow} = \lim_{x\to 0} \frac{3x^2(x^2)}{x^4+(x^2)^2} = \frac{3}{2}$$

■

Ejemplo 7. Estudiemos el límite $\lim_{(x,y)\to(0,0)} \frac{x^4y}{x^4+y^4}$. Si nos acercamos al origen por una recta del tipo $y = kx$ obtenemos

$$\lim_{(x,y)\to(0,0)} \frac{x^4y}{x^4+y^4} = \overset{y=kx}{\longrightarrow} = \lim_{x\to 0} \frac{x^4(kx)}{x^4+(kx)^4}$$
$$= \lim_{x\to 0} \frac{k}{1+k^4}x = 0$$

Esto no prueba en absoluto que el valor del límite sea cero. Lo que nos dice es que si tal límite existe, éste debe ser cero. Un argumento que concluye que el límite de hecho existe y vale cero, requiere la aplicación directa de la definición. Hagámosla. Según

$$\left|\frac{x^4y}{x^4+y^4}\right| \leq \left|\frac{x^4y}{x^4}\right| \leq |y|$$

vemos que

$$|y| \leq \sqrt{x^2+y^2} < \delta \Rightarrow \left|\frac{x^4y}{x^4+y^4}\right| \leq |y| < \delta$$

Es decir, con $\delta = \epsilon$ nos queda

$$0 < \|(x,y)-(0,0)\| < \delta \Rightarrow \left|\frac{x^4y}{x^4+y^4} - 0\right| < \epsilon$$

lo que nos dice que efectivamente $\lim_{(x,y)\to(0,0)} \frac{x^4y}{x^4+y^4} = 0$. Continuando con este mismo ejemplo, veamos cómo se pueden usar las coordenadas polares para convencerse de que tal límite existe y vale cero. Decir que $(x, y) \to (0, 0)$ equivale a decir que (en coordenadas polares) $r \to 0$ (independientemente del valor de θ). Expresando la función $f(x, y) = \frac{x^4y}{x^4+y^4}$ en cordenadas polares $x = r\cos\theta$, $y = r \operatorname{sen}\theta$ obtenemos la función

$$g(r,\theta) = \frac{\cos^4\theta \operatorname{sen}\theta}{\cos^4\theta + \operatorname{sen}^4\theta} r$$

Obsérvese que ésta es el producto de la función $\varphi(\theta) = \frac{\cos^4\theta \operatorname{sen}\theta}{\cos^4\theta+\operatorname{sen}^4\theta}$ que está acotada por la función $\psi(r) = r$ que tiende a cero cuando $r \to 0$. Podemos concluir entonces que

$$\lim_{r\to 0} g(r,\theta) = 0$$

Hemos usado el siguiente resultado para que el lector lo demuestre: "si $\varphi(\theta)$ es una función acotada (en alguna bola con centro en el origen) y $\psi(r)$ es una función que tiende a cero cuando r tiende a cero, entonces $\lim_{r\to 0} g(r,\theta) = \varphi(\theta)\psi(r) = 0$". ■

En el siguiente teorema se recogen las principales propiedades de los límites de funciones de varias variables. Como el lector podrá advertir, éstas son completamente análogas a aquéllas que conocía para funciones de una sola variable (lo cual no debe ser en absoluto extraño, pues el concepto de límite es el mismo para funciones de una o varias variables).

Teorema 2.3.1 Sean $f, g\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones definidas en el abierto U de $\mathbb{R}^n$ y sea $\mathbf{x}_0$ un punto de U o un punto de la frontera de U. Suponga que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = L \qquad \text{y} \qquad \lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = M$$

Entonces:

1. $\lim_{\mathbf{x}\to\mathbf{x}_0}(f+g)(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{x}_0}(f(\mathbf{x})+g(\mathbf{x})) = \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) + \lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = L + M$
2. $\lim_{\mathbf{x}\to\mathbf{x}_0}(fg)(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})g(\mathbf{x}) = \left(\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})\right)\left(\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x})\right) = LM$
3. Si $M \neq 0$, entonces $\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f}{g}(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})}{\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x})} = \frac{L}{M}$ ■

Demostración (de 1). (Opcional) Como $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = L$, tenemos que dado $\epsilon > 0$, existe $\delta_1 > 0$ tal que

$$\mathbf{x} \in B(\mathbf{x}_0, \delta_1) \cap U,\ \mathbf{x} \neq \mathbf{x}_0 \Rightarrow f(\mathbf{x}) \in B(L, \epsilon/2)$$

Del mismo modo, como $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = M$, se tiene que (dado $\epsilon > 0$) existe $\delta_2 > 0$ tal que

$$\mathbf{x} \in B(\mathbf{x}_0, \delta_2) \cap U,\ \mathbf{x} \neq \mathbf{x}_0 \Rightarrow g(\mathbf{x}) \in B(M, \epsilon/2)$$

Se quiere ver que dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$\mathbf{x} \in B(\mathbf{x}_0, \delta) \cap U,\ \mathbf{x} \neq \mathbf{x}_0 \Rightarrow f(\mathbf{x}) + g(\mathbf{x}) \in B(L + M, \epsilon)$$

Tome $\delta = \min(\delta_1, \delta_2)$. En tal caso

$$\mathbf{x} \in B(\mathbf{x}_0, \delta) \cap U,\ \mathbf{x} \neq \mathbf{x}_0 \ \Rightarrow\ \mathbf{x} \in B(\mathbf{x}_0, \delta_1) \cap B(\mathbf{x}_0, \delta_2) \cap U,\ \mathbf{x} \neq \mathbf{x}_0$$

de modo que al aplicar la desigualdad triangular tenemos (para $\mathbf{x} \in B(\mathbf{x}, \delta) \bigcap U, \mathbf{x} \neq \mathbf{x}_0$)

$$\begin{aligned} |f(\mathbf{x}) + g(\mathbf{x}) - (L+M)| &= |(f(\mathbf{x}) - L) + (g(\mathbf{x}) - M)| \\ &\leq |f(\mathbf{x}) - L| + |g(\mathbf{x}) - M| < \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

es decir $f(\mathbf{x}) + g(\mathbf{x}) \in B(L + M, \epsilon)$, como queríamos. Q.E.D.

Un corolario del teorema anterior y del ejemplo 2 es el siguiente (su demostración la dejamos como ejercicio para el lector):

Corolario Si $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es una función polinomial, es decir, una función del tipo

$$f(x, y) = \sum_{i,j} a_{ij} x^i y^j$$

donde a_{ij} son constantes dadas, entonces

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$$ ■

Pasemos ahora a estudiar el concepto de continuidad para funciones de varias variables.

De manera intuitiva, debemos pensar en la continuidad de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto $\mathbf{x}_0 \in U$ como la propiedad que nos dice que los valores de $f(\mathbf{x})$ "cambian suavemente" cuando $\mathbf{x}$ se mueve "cerca de $\mathbf{x}_0$". Es decir, para que la función f sea continua en $\mathbf{x}_0$, si $\mathbf{x}$ es una vecina íntima de $\mathbf{x}_0$, se debe tener que $f(\mathbf{x})$ es una vecina íntima de $f(\mathbf{x}_0)$ (¡valor que debe existir!). Más aún, se tiene la definición siguiente.

Definición. (continuidad en un punto). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el abierto U de $\mathbb{R}^n$ y sea $\mathbf{x}_0 \in$ U. Se dice que f es una función *continua* en $\mathbf{x}_0$ si

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Si la función f no es continua en $\mathbf{x}_0$, se dice que es *discontinua* en ese punto. ■

Veamos algunos ejemplos.

Ejemplo 8. Una función polinomial $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es continua en cualquier punto $(x_0, y_0) \in \mathbb{R}^2$. Esto es consecuencia del corolario del teorema 2.3.1. ■

Ejemplo 9. La función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, definida como

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

es discontinua en (0, 0), pues se vió (ejemplo 5) que el límite

$$\lim_{(x,y)\to(0,0)} f(x, y)$$

no existe. ■

Ejemplo 10. La función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, definida como

$$f(x, y) = \begin{cases} \dfrac{x^3 y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

es continua en cualquier punto $(x_0, y_0) \in \mathbb{R}^2$. En efecto, si $(x_0, y_0) \neq (0, 0)$, al seguir el teorema 2.3.1 vemos que

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = \lim_{(x,y)\to(x_0,y_0)} \frac{x^3 y}{x^2 + y^2} = \frac{x_0^3 y_0}{x_0^2 + y_0^2} = f(x_0, y_0)$$

en tanto que si $(x_0, y_0) = (0, 0)$, aplicando la técnica con coordenadas polares mencionada en el ejemplo 7 tenemos

$$\lim_{(x,y)\to(0,0)} \frac{x^3 y}{x^2 + y^2} = \lim_{r\to 0} \frac{(r\cos\theta)^3 (r \operatorname{sen}\theta)}{(r\cos\theta)^2 + (r\operatorname{sen}\theta)^2} = \lim_{r\to 0} (\cos^3\theta \operatorname{sen}\theta) r^2 = 0 = f(0, 0).$$ ■

Definición. (continuidad en un abierto) Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el abierto U de $\mathbb{R}^n$. Se dice que f es continua en U (o simplemente que f es continua) si lo es para todos y cada uno de los puntos $(x, y) \in U$. ■

Ejemplo 11. Una función polinomial (ver ejemplo 8) es continua en todo $\mathbb{R}^2$. La función del ejemplo 10 es continua en todo $\mathbb{R}^2$. La función $f(x, y) = \frac{xy}{x^2+y^2}$ es continua en $\mathbb{R}^2 - \{(0, 0)\}$ (i.e. es continua en todo su dominio). ■

Al poner junto la definición de continuidad y el teorema 2.3.1, se obtiene fácilmente el siguiente teorema

Teorema 2.3.2 Sean $f, g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ funciones definidas en el abierto U de $\mathbb{R}^n$. Si f y g son continuas, entonces

1. la función $f + g: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ es continua.
2. la función $fg: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ es continua.
3. la función $\frac{f}{g}: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $\left(\frac{f}{g}\right)(\mathbf{x}) = \frac{f(\mathbf{x})}{g(\mathbf{x})}$ es continua en todo punto $\mathbf{x} \in U$ donde $g(\mathbf{x}) \neq 0$.

Es decir, suma y producto de funciones continuas es una función continua. El cociente también lo es en aquellos puntos en que no se anula el denominador. ■

Por último, considere la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ y $g: I \subseteq \mathbb{R} \to \mathbb{R}$ en donde I es un intervalo de $\mathbb{R}$ que contiene las imágenes de f (es decir, se tiene que rango de $f \subseteq I$). Podemos formar la composición $g \circ f: U \subseteq \mathbb{R}^n \to \mathbb{R}$
Supongamos que f es continua en $\mathbf{x}_0 \in U$ y que g es continua en $f(\mathbf{x}_0) \in I$. Un razonamiento intuitivo nos dice que la composición $g \circ f$ debe ser continua en $\mathbf{x}_0$. En efecto: veamos que si $\mathbf{x}$ está cerca de $\mathbf{x}_0$ entonces $(g \circ f)(\mathbf{x})$ debe estar cerca de $(g \circ f)(\mathbf{x}_0)$. Siendo $\mathbf{x}$ vecina íntima de $\mathbf{x}_0$ (en U), la imagen $f(\mathbf{x})$ será vecina íntima de $f(\mathbf{x}_0)$ (en I), pues f es continua en $\mathbf{x}_0$. Ahora, si $f(\mathbf{x})$ es vecina íntima de $f(\mathbf{x}_0)$, la imagen $g(f(\mathbf{x}))$ será vecina íntima de $g(f(\mathbf{x}_0))$, debido a la continuidad de g. Poniendo todo junto, tenemos: $\mathbf{x}$ cerca de $\mathbf{x}_0$ produce $(g \circ f)(\mathbf{x})$ cerca de $(g \circ f)(\mathbf{x}_0)$. Esto nos dice que la composición $g \circ f$ es continua en $\mathbf{x}_0$. Con algunos épsilons (ϵ) y deltas (δ) podemos convencernos "rigurosamente" de la validez de este argumento. Esto lo dejamos para el lector.

En resumen, añadiendo a lo establecido en el teorema 2.3.2 tenemos también que "composición de funciones continuas es continua".

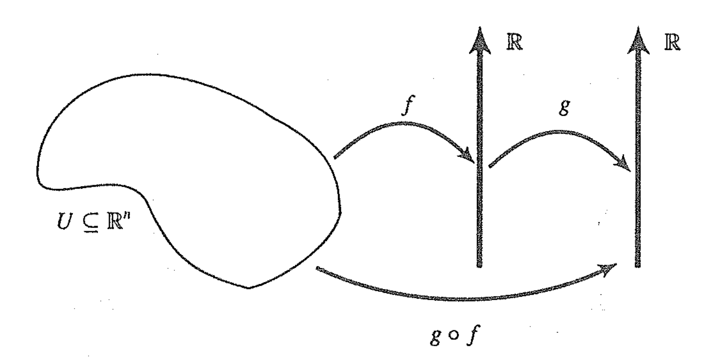


Figura 10. Composición de las funciones g con f.

Ejemplo 12. La función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = \operatorname{sen}\left(\frac{x^2+y^2}{5x^4+y^4+2}\right)$ es continua en $\mathbb{R}^2$, pues la composición de la función $f(x, y) = \frac{x^2+y^2}{5x^4+y^4+2}$ que es continua, con la función $g(x) = \operatorname{sen} x$, que también lo es. ■

Ejemplo 13. La función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$,

$$F(x, y) = e^{(x^2+y^2)} + \left|\cos \frac{x+y}{x^2+1}\right| + \frac{e^{xy}\operatorname{sen}(x+y)}{1+e^{\sqrt{x^2+y^2}}}$$

es continua en $\mathbb{R}^2$ pues está formada por sumas, productos y cocientes de funciones que son composiciones de funciones continuas y que, por lo tanto, son continuas. ■

Ejercicios (Capítulo 2, Sección 3)

1. Escriba explícitamente (como conjunto de puntos en el espacio $\mathbb{R}^n$ correspondiente) cada una de las siguientes bolas abiertas.
   a. $B(3, 0.5)$ en $\mathbb{R}$.
   b. $B((2, -3), 0.1)$ en $\mathbb{R}^2$.
   c. $B((1, 1, 4), 2)$ en $\mathbb{R}^3$.
   d. $B((2, -1, 9, 3, 5), 1)$ en $\mathbb{R}^5$.
2. Verifique que el semiplano superior $A = \{(x, y) | y > 0\}$ es un conjunto abierto, tomando un punto cualquiera $\mathbf{p} = (x_0, y_0) \in A$ y consiguiendo una bola B con centro en $\mathbf{p}$ y (algún) radio $r > 0$ que quede completamente contenida en A.
3. Demuestre que el conjunto vacío $\varnothing$ es un conjunto abierto en $\mathbb{R}^n$. (Sugerencia: si $\varnothing$ no fuera abierto...)
4. Demuestre que el espacio $\mathbb{R}^n$ es un conjunto abierto en $\mathbb{R}^n$.
5. ¿Verdadero o falso? La intersección de dos conjuntos abiertos en $\mathbb{R}^n$ o es vacía o contiene una infinidad de puntos (es decir, no puede constar de un número finito de puntos). Justifique su respuesta.

6. Sean A y B dos conjuntos abiertos en $\mathbb{R}^n$. Demuestre que su unión $A \cup B$ y su intersección $A \cap B$ son también conjuntos abiertos en $\mathbb{R}^n$. Más en general, sea $A_1, A_2, \ldots, A_n$ una colección (*finita*) de conjuntos abiertos en $\mathbb{R}^n$. Demuestre que su unión $\bigcup_{i=1}^{n} A_i$ y su intersección $\bigcap_{i=1}^{n} A_i$ son conjuntos abiertos en $\mathbb{R}^n$.

7. Demuestre que el conjunto $A = \{\mathbf{p}\}$ formado por un solo punto $\mathbf{p} \in \mathbb{R}^n$ *no* es un conjunto abierto en $\mathbb{R}^n$.

8. En este ejercicio se establece el importante resultado sobre conjuntos abiertos en $\mathbb{R}^n$: "la unión arbitraria de conjuntos abiertos es un conjunto abierto y la intersección *finita* de conjuntos abiertos es un conjunto abierto" (es decir, el resultado establecido en el ejercicio 6 se puede generalizar a una colección arbitraria —no necesariamente finita— de conjuntos abiertos, solamente en el caso de la unión). Sea $A_1, A_2, \ldots, A_n, \ldots$ una colección arbitraria de conjuntos abiertos en $\mathbb{R}^n$.

   a. Demuestre que el conjunto $\bigcup_i A_i$ (unión de los conjuntos A_i) es abierto.

   b. Demuestre que el conjunto $\bigcap_i A_i$ (intersección de los conjuntos A_i) puede no ser un conjunto abierto. (Sugerencia: tome $A_i = B(\mathbf{p}, \frac{1}{i})$ = bola abierta con centro en $\mathbf{p} \in \mathbb{R}^n$ y radio $\frac{1}{i}$, $i = 1, 2, \ldots$ Se sabe que cada A_i es un conjunto abierto. Verifique que $\bigcap_i A_i = \{\mathbf{p}\}$. Concluya...).

9. (Conjuntos cerrados de $\mathbb{R}^n$). Se dice que un conjunto $C \subseteq \mathbb{R}^n$ es *cerrado* (en $\mathbb{R}^n$) si su complemento $\mathbb{R}^n - C$ es abierto.

   a. Demuestre que todo el espacio $\mathbb{R}^n$ y el conjunto vacío $\varnothing$ son conjuntos cerrados. (Moraleja: ser cerrado *no* es lo contrario de ser abierto. Hay conjuntos que son cerrados y abiertos a la vez, como $\mathbb{R}^n$ y $\varnothing$. Y, por supuesto, hay conjuntos que no son ni cerrados ni abiertos).

   b. Demuestre que los intervalos cerrados de $\mathbb{R}$ son conjuntos cerrados.

   c. Defina la *bola cerrada* (en $\mathbb{R}^n$) con centro en $\mathbf{p}$ (un punto dado en $\mathbb{R}^n$) y radio $r > 0$, denotada por $\overline{B(\mathbf{p}, r)}$, como $\overline{B(\mathbf{p}, r)} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{p}\| \leq r\}$. Obsérvese que $\overline{B(\mathbf{p}, r)}$ es la unión de la bola abierta $B(\mathbf{p}, r)$ con su frontera (la esfera $\|\mathbf{x} - \mathbf{p}\| = r$). Demuestre que es $\overline{B(\mathbf{p}, r)}$ un conjunto cerrado.

   d. Sea A un conjunto cualquiera de $\mathbb{R}^n$. Demuestre que su frontera ∂A es un conjunto cerrado.

   e. Demuestre que la unión finita de conjuntos cerrados es un conjunto cerrado. Por medio de un ejemplo muestre que esto puede ser falso si se consideran uniones arbitrarias de cerrados. (Sugerencia: use el resultado del inciso **b.** del ejercicio anterior. Recuerde que $\mathbb{R}^n - \bigcap_i A_i = \bigcup_i (\mathbb{R}^n - A_i)$).

   f. Demuestre que la intersección arbitraria de conjuntos cerrados es un conjunto cerrado.

En cada uno de los ejercicios 10–26, se da un conjunto A de puntos en el plano $\mathbb{R}^2$. Diga en cada caso si el conjunto dado es abierto, cerrado (en $\mathbb{R}^2$), o no es ni abierto ni cerrado, justificando su respuesta. Haga una gráfica que muestre el conjunto dado en el plano xy.

10. $A = \{(x, y) | xy > 0\}$

11. $A = \{(x, y) | x > 0, y > 0, y < 2 - x\}$

12. $A = \{(x, y) | |x| + |y| < 1\}$

13. $A = \{(x, y) | y = x\}$

14. $A = \{(x, y)|x^2 + y^2 > 1\}$

15. $A = \{(x, y)|x^2 + y^2 \leq 1\}$

16. $A = \{(x, y)|x^2 + y^2 \leq -1\}$

17. $A = \{(x, y)|y < x^2\}$

18. $A = \{(x, y)|y = x^2\}$

19. $A = \{(x, y)|2 < x < 4, 2 < y < 5\}$

20. $A = \{(x, y)|2 < x < 4, 2 < y \leq 5\}$

21. $A = \{(x, y)|\operatorname{sen} x < 0\}$

22. $A = \{(x, y)|\cos y \geq 0\}$

23. $A = \{(x, y)|x^2 + y^2 \geq 0\}$

24. $A = \{(x, y)|(x^2 + y^2 - 4)(1 - x^2 - y^2) > 0\}$

25. $A = \{(x, y)|(y - x^2 + 2)(y + x^2 - 4) < 0\}$

26. $A = \{(x, y)|(x^2 + y^2 - 1)(2x - x^2 - y^2) > 0\}$

En cada uno de los ejercicio 27–30, diga si el conjunto A de $\mathbb{R}^n$ dado es abierto, cerrado, o ninguno de los dos, justificando su respuesta.

27. $A = \{(x_1, x_2, \ldots, x_n)|x_i > 0, i = 1, 2, \ldots, n\}$

28. $A = \{(x_1, x_2, \ldots, x_n)|x_1 x_2 \ldots x_n < 0\}$

29. $A = \{(x_1, x_2, \ldots, x_n)|x_1 \leq 0\}$

30. $A = \{(x_1, x_2, \ldots, x_n)|x_1^2 + x_2^2 + \cdots + x_n^2 \geq 1\}$

31. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = 4y$. Se sabe que

$$\lim_{(x,y)\to(1,2)} f(x, y) = 8$$

Dado $\epsilon = 0.1$, halle $\delta > 0$ tal que

$$\|(x, y) - (1, 2)\| < \delta \Rightarrow |f(x, y) - 8| < \epsilon$$

32. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = -3x$. Se sabe que

$$\lim_{(x,y)\to(3,7)} f(x, y) = -9$$

Dado $\epsilon = 0.4$, halle $\delta > 0$ tal que

$$\|(x, y) - (3, 7)\| < \delta \Rightarrow |f(x, y) + 9| < \epsilon$$

33. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = 5x - 2y$. Se sabe que

$$\lim_{(x,y)\to(1,1)} f(x, y) = 3$$

Dado $\epsilon = 0.2$, halle $\delta > 0$ tal que

$$\|(x, y) - (1, 1)\| < \delta \Rightarrow |f(x, y) - 3| < \epsilon.$$

34. Sea $f: \mathbb{R}^n \to \mathbb{R}$ la función $f(x_1, x_2, \ldots, x_n) = ax_i$ en donde a es un real dado no nulo e i es un índice fijo, $1 \leq i \leq n$. Sea $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ un punto dado de $\mathbb{R}^n$. Demuestre que $\lim_{\mathbf{x} \to \bar{\mathbf{x}}} f(\mathbf{x}) = a\bar{x}_i$.

Para cada una de las funciones $z = f(x, y)$ dadas en los ejercicios 35–38:

a. Diga dónde están definidas.

b. Demuestre que los límites

$$\lim_{x \to 0}\left(\lim_{y \to 0} f(x, y)\right) \qquad \text{y} \qquad \lim_{y \to 0}\left(\lim_{x \to 0} f(x, y)\right)$$

(llamados límites iterados) existen y valen cero. ¿Cómo estamos haciendo tender el punto (x, y) al origen al hacer el cálculo de estos límites? ¿Puede concluir de aquí que el límite

$$\lim_{(x,y) \to (0,0)} f(x, y)$$

existe y vale 0?

c. Demuestre que $\lim_{(x,y) \to (0,0)} f(x, y) = 0$ si el punto (x, y) se acerca a $(0, 0)$ por rectas del tipo $y = kx$. ¿Puede concluír de aquí que tal límite existe y vale 0?

d. Demuestre que $\lim_{(x,y) \to (0,0)} f(x, y) = 0$ si el punto (x, y) se acerca a (0,0) por parábolas del tipo $y = kx^2$. ¿Puede concluír de aquí que tal límite existe y vale 0?

e. Use la definición de límite para demostrar que el límite de $f(x, y)$ cuando (x, y) tiende a $(0, 0)$ efectivamente existe y vale 0.

f. Use coordenadas polares para concluir nuevamente que el límite de $f(x, y)$ cuando (x, y) tiende a (0,0) existe y vale 0.

35. $f(x, y) = \dfrac{y^3}{x^2 + y^2}$.

36. $f(x, y) = \dfrac{3x^3 y^2}{x^2 + y^2}$.

37. $f(x, y) = \dfrac{7x^2 y^2}{2x^2 + 2y^2}$.

38. $f(x, y) = \dfrac{x^3 y^4}{x^4 + y^4}$.

Para cada una de las funciones $z = f(x, y)$ dadas en los ejercicio 39–46, demuestre que el límite $\lim_{(x,y) \to (0,0)} f(x, y)$ no existe.

39. $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$.

40. $f(x, y) = \dfrac{x^2 y}{x^3 + y^3}$.

41. $f(x, y) = \dfrac{xy^2}{y^4 + x^2}$.

42. $f(x, y) = \dfrac{2xy^4}{x^5 + 6y^5}$.

43. $f(x, y) = \dfrac{x^3 y^2}{x^6 + y^4}$.

44. $f(x, y) = \dfrac{x^4 y}{x^8 + y^2}$.

45. $f(x, y) = \dfrac{8x^3 y^2}{x^9 + y^3}$.

46. $f(x, y) = \dfrac{y^3 x}{y^6 + x^2}$.

47. Demuestre que una condición *necesaria* para que el límite $\text{lím}_{(x,y)\to(x_0,y_0)} f(x, y)$ exista y sea L, es que si los límites (iterados)

$$\lim_{x\to x_0}\left(\lim_{y\to y_0} f(x, y)\right) \qquad \text{y} \qquad \lim_{y\to y_0}\left(\lim_{x\to x_0} f(x, y)\right)$$

existen, deben valer L.

a. ¿Cuál es la negación de la afirmación anterior?

b. Use la función

$$f(x, y) = \frac{x^2 y^2}{x^2 y^2 + (x - y)^2}$$

para demostrar que tal condición no es suficiente.

c. Para la función $f(x, y) = (x + y) \operatorname{sen} \frac{1}{x} \operatorname{sen} \frac{1}{y}$, demuestre que los límites iterados no existen, pero que el límite $\text{lím}_{(x,y)\to(0,0)} f(x, y)$ sí existe. ¿Contradice este ejemplo el resultado general establecido en este ejercicio?

48. Considere la función $f(x, y, z) = \dfrac{x + y + z}{x + y - z}$. ¿Dónde está definida? Demuestre que el límite $\text{lím}_{(x,y,z)\to(0,0,0)} f(x, y, z)$ no existe. (Sugerencia: haga tender (x, y, z) a $(0, 0, 0)$ por los ejes coordenados).

49. Sea $f(x, y, z) = \dfrac{2x^2 + y^2 - z^2}{x^2 - y^2}$. ¿Dónde está definida esta función? Demuestre que el límite $\text{lím}_{(x,y,z)\to(0,0,0)} f(x, y, z)$ no existe.

50. Sea $f(x, y, z) = \dfrac{xyz}{x^3 + y^3 + z^3}$. ¿Dónde está definida esta función? Demuestre que el límite $\text{lím}_{(x,y,z)\to(0,0,0)} f(x, y, z)$ no existe. (Sugerencia: acérquese al origen por los ejes coordenados y por la recta $x = y = z$).

51. Sea $f(x, y, z) = \dfrac{x^2 z^3 y}{x^6 + z^6}$. ¿Dónde está definida esta función? Demuestre que el límite $\text{lím}_{(x,y,z)\to(0,0,0)} f(x, y, z)$ no existe.

52. Sea $f: I \subseteq \mathbb{R} \to \mathbb{R}$ una función definida en el conjunto I de $\mathbb{R}$. Sea x_0 un punto de I o un punto frontera de I. Suponga que $\lim_{x \to x_0} f(x) = L$. Considere la función $\tilde{f}: I \times \mathbb{R} \subseteq \mathbb{R}^2 \to \mathbb{R}$ dada por $\tilde{f}(x, y) = f(x)$. ¿Cómo es la gráfica de la función $\tilde{f}$? Demuestre que el límite $\lim_{(x,y) \to (x_0, y_0)} \tilde{f}(x, y)$ existe y vale L (en donde y_0 es un número real cualquiera).

53. Sean $f, g: I \subseteq \mathbb{R} \to \mathbb{R}$ dos funciones definidas en el conjunto I de $\mathbb{R}$. Sean x_0, y_0 dos puntos de I o puntos frontera de I. Suponga que

$$\lim_{x \to x_0} f(x) = L \qquad \text{y} \qquad \lim_{y \to y_0} g(y) = M$$

a. Considere la función $\tilde{f}(x, y) = f(x) + g(y)$. ¿En dónde está definida $\tilde{f}$? Demuestre que el límite $\lim_{(x,y) \to (x_0, y_0)} \tilde{f}(x, y)$ existe y vale $L + M$.

b. Considere la función $\tilde{f}(x, y) = f(x)g(y)$. ¿En dónde está definida $\tilde{f}$? Demuestre que el límite $\lim_{(x,y) \to (x_0, y_0)} \tilde{f}(x, y)$ existe y vale LM.

c. Suponga que $M \neq 0$. Considere la función $\tilde{f}(x, y) = f(x)/g(y)$. ¿En dónde está definida $\tilde{f}$? Demuestre que el límite $\lim_{(x,y) \to (x_0, y_0)} \tilde{f}(x, y)$ existe y vale L/M.

En los ejercicios 54–61, calcule los límites indicados.

54. $\lim_{(x,y) \to (1,1)} \left[\frac{x^2 - 1}{x - 1} + \frac{y - 1}{y^2 - 1} \right]$

55. $\lim_{(x,y) \to (1,1)} \frac{(x^3 - 1)(y^4 - 1)}{(x - 1)(y^2 - 1)}$

56. $\lim_{(x,y) \to (0,0)} \frac{\operatorname{sen} x \operatorname{sen} 3y}{2xy}$

57. $\lim_{(x,y) \to (0,1)} \frac{(y^2 + 2y - 3)(1 - \cos x)}{x^2(y - 1)}$

58. $\lim_{(x,y) \to (0,0)} \frac{(1 - \cos 2x)(\cos 3y - 1)}{5x^2 y}$

59. $\lim_{(x,y) \to (0,0)} \frac{\operatorname{arcsen}(2x) \arctan(3y)}{xy}$

60. $\lim_{(x,y) \to (1,2)} \frac{(x^3 + x^2 - 5x + 3)(y^2 - 4y + 4)}{(y^4 - 4y^3 + 7y^2 - 12y + 12)(x^3 - 4x^2 + 5x - 2)}$

61. $\lim_{(x,y) \to (0,0)} \frac{(e^x - 1)(e^{2y} - 1)}{xy}$

62. (Límites al infinito). Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función definida en todo el espacio $\mathbb{R}^2$ (excepto posiblemente en un subconjunto acotado). Se dice que el límite de f cuando x, y tienden a infinito, es L, lo cual se escribe como

$$\lim_{\substack{x \to \infty \\ y \to \infty}} f(x, y) = L$$

si dado $\epsilon > 0$ existe $N > 0$ tal que

$$\|(x, y)\| > N \Rightarrow |f(x, y) - L| < \epsilon$$

a. Demuestre que $\lim\limits_{\substack{x\to\infty \\ y\to\infty}} \dfrac{1}{x+y} = 0.$

b. Demuestre que $\lim\limits_{\substack{x\to\infty \\ y\to\infty}} e^{-(x^2+y^2)} = 0.$

c. Sea $z = \phi(y)$ una función (cuya gráfica está en el plano zy de $\mathbb{R}^3$) tal que $\lim_{y\to\infty} \phi(y) = L$. Sea $z = \phi\left(\sqrt{x^2+y^2}\right)$ la superficie de revolución que se obtiene al girar $z = \phi(y)$ alrededor del eje z. Demuestre que

$$\lim_{\substack{x\to\infty \\ y\to\infty}} \phi\left(\sqrt{x^2+y^2}\right) = L$$

En los ejercicios 63–72, diga en dónde la función dada es continua, justificando en cada caso su respuesta con los resultados generales sobre continuidad discutidos en esta sección

63. $f(x, y) = x^2 + 4xy + 5y^2 - 7x + 9y - 10$

64. $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$

65. $f(x, y) = \dfrac{x^2 + y^2}{x^2 - y^2}$

66. $f(x, y) = \dfrac{2x + 3y^5}{x^2 + y^2 + 1}$

67. $f(x, y) = \operatorname{sen} x + \operatorname{sen} y$

68. $f(x, y) = \dfrac{\operatorname{sen} x}{\operatorname{sen} y}$

69. $f(x, y) = \operatorname{sen}^2(x^3 \cos^4 y)$

70. $f(x, y) = \begin{cases} \dfrac{x^3y^2}{x^4 + 3y^4} & \text{si } (x, y) \neq (0, 0) \\ 1 & \text{si } (x, y) = (0, 0) \end{cases}$

71. $f(x, y) = \begin{cases} \dfrac{x^4 - 3y^4}{x^4 + 5y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$

72. $f(x, y) = \begin{cases} \dfrac{6x^3y^3}{x^4 + 7y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$

En los ejercicios 73–77 se da una función $z = f(x, y)$ que no está definida en $(0, 0)$. ¿Es posible definir el valor $f(0, 0)$ de tal modo que f sea continua en este punto? Explique.

73. $f(x, y) = \dfrac{3x^2y}{x^4 + y^4}$

74. $f(x, y) = \dfrac{3x^2y^3}{x^4 + y^4}$

75. $f(x, y) = \dfrac{5x^2y^2}{x^3 + y^6}$

76. $f(x, y) = \dfrac{3x^2y^8}{x^8 + y^8}$

77. $f(x, y) = \dfrac{x - y}{x + y}$

78. Estudie la continuidad de la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \operatorname{sgn}(xy)$.

79. En el texto se estableció que "suma y producto de funciones continuas es una función continua" (es decir, si f y g son funciones continuas, entonces $f + g$ y fg son continuas). Por medio de un ejemplo concreto, muestre que la afirmación recíproca es falsa.

80. Se dice que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, definida en el conjunto abierto U de $\mathbb{R}^2$, es continua *respecto de su primera variable* (*respecto de* su *segunda variable*) en el punto $(x_0, y_0) \in U$, si la función $\phi(x) = f(x, y_0)$ es continua en x_0 (si la función $\psi(y) = f(x_0, y)$ es continua en y_0, respectivamente).

a. Demuestre que f es continua respecto de su primera variable en (x_0, y_0) si y sólo si $\lim_{h \to 0} f(x_0 + h, y_0) = f(x_0, y_0)$.

b. Demuestre que f es continua respecto de su segunda variable en (x_0, y_0) si y sólo si $\lim_{h \to 0} f(x_0, y_0 + h) = f(x_0, y_0)$.

c. Dé un ejemplo de una función f (y un punto $\mathbf{p} = (x_0, y_0)$ de su dominio) tal que:

c1. f sea continua respecto de su primera variable y discontinua respecto de su segunda variable en $\mathbf{p}$.

c2. f sea continua respecto de su primera y segunda variable en $\mathbf{p}$.

d. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

es continua respecto de su primera y segunda variables en $(0, 0)$.

e. Demuestre que si f es continua en (x_0, y_0), entonces es continua respecto de su primera y segunda variables en ese punto.

f. Demuestre que la afirmación recíproca del inciso anterior es falsa, considerando la función del ejemplo d).

81. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$. Sea $\mathbf{v} \in \mathbb{R}^2$ un vector no nulo dado. Se dice que f es *continua* en la *dirección del vector* $\mathbf{v}$, en el punto $\mathbf{p} = (x_0, y_0) \in U$, si

$$\lim_{t \to 0} f(\mathbf{p} + t\mathbf{v}) = f(\mathbf{p})$$

a. Demuestre que f es continua en la dirección del vector $\mathbf{v} = i = (1, 0)$ en $\mathbf{p}$ si y sólo si f es continua respecto de su primera variable en $\mathbf{p}$.

b. Demuestre que f es continua en la dirección del vector $\mathbf{v} = j = (0, 1)$ en $\mathbf{p}$ si y sólo si f es continua respecto de su segunda variable en $\mathbf{p}$.

c. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

es continua en la dirección de cualquier vector no nulo $\mathbf{v} \in \mathbb{R}^2$ en el punto $(0, 0)$.

d. Demuestre que si f es continua en (x_0, y_0) entonces f es continua en la dirección de cualquier vector no nulo $\mathbf{v} \in \mathbb{R}^2$ en ese punto.

e. Demuestre que la afirmación recíproca del inciso anterior es falsa, considerando la función del inciso **c**.

82. Demuestre que la función $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|\mathbf{x}\|$ es continua. (Sugerencia: use la desigualdad establecida en el ejercicio 8 de la sección 3 del capítulo 1).

(*) **83.** Demuestre que la función $f: \mathbb{R}^n \to \mathbb{R}$ es continua en el punto $\mathbf{p}$ si y sólo si la imagen inversa de cualquier conjunto abierto (en $\mathbb{R}$) que contenga a $f(\mathbf{p})$ es un conjunto abierto (en $\mathbb{R}^n$) que contiene a $\mathbf{p}$. Es decir, si llamamos $I_{f(\mathbf{p})}$ a un conjunto abierto de $\mathbb{R}$ que contiene a $f(\mathbf{p})$, entonces (demuestre que) f es continua en $\mathbf{p}$ si y sólo si el conjunto

$$f^{-1}(I_{f(\mathbf{p})}) = \{x \in \mathbb{R}^n | f(x) \in I_{f(p)}\} \subseteq \mathbb{R}^n$$

es abierto.

2.4 Derivadas parciales

En esta sección estudiaremos el primer concepto "diferenciable" para funciones de varias variables. Recordemos que para una función de una variable $f: I \subseteq \mathbb{R} \to \mathbb{R}$ definida en el intervalo abierto I de $\mathbb{R}$, se define la *derivada* de f en $x_0 \in I$, denotada por $f'(x_0)$, como el valor del límite

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

cuando éste existe (en cuyo caso decimos que f es *diferenciable* en x_0). Si $f'(x_0)$ existe, su valor nos da la pendiente de la recta tangente a la gráfica de la función $y = f(x)$ en el punto $(x_0, f(x_0))$.

Las funciones importantes a estudiar, bajo la óptica del cálculo, son las funciones diferenciables. Cuando se tiene una función diferenciable, lo importante es obtener información de la función a partir de su derivada (la información que se obtiene de f a partir del valor de $f'(x_0)$ es *local*, alrededor de x_0). Por ejemplo, el simple hecho de la existencia de $f'(x_0)$ nos habla del comportamiento suave de la gráfica de la función en los alrededores del punto $(x_0, f(x_0))$; el signo de $f'(x_0)$ nos habla del crecimiento y/o decrecimiento de la función alrededor del punto, etc. Este es, en realidad, uno de los objetivos principales del cálculo: obtener información de una función diferenciable a partir de su derivada.

Resulta deseable, por tanto, disponer de un concepto de "diferenciabilidad" para funciones de varias variables semejante a aquél que conocemos para funciones de una variable. En esta sección estudiaremos un primer acercamiento a este importante concepto, el cual será abordado detalladamente en la sección 6 de este capítulo.

Comencemos por considerar una función de dos variables $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el abierto U de $\mathbb{R}^2$. Sea $\mathbf{p} = (x_0, y_0)$ un punto de U. Se define la derivada parcial de f con respecto de x (la primera variable de f) en el punto $\mathbf{p}$, denotada por $\frac{\partial f}{\partial x}(\mathbf{p})$ (o $f_x(\mathbf{p})$, o $f_1(\mathbf{p})$, o $D_1 f(\mathbf{p})$), como el límite

$$\frac{\partial f}{\partial x}(\mathbf{p}) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

cuando éste límite existe. Del mismo modo, la derivada parcial de f con respecto a y (la segunda variable de f) en $\mathbf{p}$, denotada por $\frac{\partial f}{\partial y}(\mathbf{p})$ (o $f_y(\mathbf{p})$, o $f_2(\mathbf{p})$, o $D_2 f(\mathbf{p})$), es el límite (si existe)

$$\frac{\partial f}{\partial y}(\mathbf{p}) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

Observando con detenimiento esta definición, vemos que tiene los mismos ingredientes de la definición de derivada de una función de una sola variable. Esto es: se toma la función con "su" variable incrementada en h, se le resta el valor de la función en el punto en cuestión, se divide todo entre el incremento, y se toma el límite cuando éste tiende a cero. En el caso de una función de una sola variable se hace este proceso con *su* variable. En el caso de una función con dos variables este proceso se efectúa *respecto de cada variable* conservando fija la otra.

De otro modo, si consideramos la función $\varphi(x) = f(x, y_0)$ que se obtiene de *fijar* en la función $f(x, y)$ la segunda variable en $y = y_0$, sabemos que la derivada de φ en $x = x_0$ es

$$\varphi'(x_0) = \lim_{h \to 0} \frac{\varphi(x_0 + h) - \varphi(x_0)}{h} = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \frac{\partial f}{\partial x}(\mathbf{p})$$

Similarmente se tiene que

$$\psi'(y_0) = \frac{\partial f}{\partial y}(\mathbf{p})$$

en donde $\psi(y) = f(x_0, y)$.

Obsérvese que desde el punto de vista geométrico, la función $\varphi(x) = f(x, y_0)$ representa la curva que se obtiene al intersectar la superficie $z = f(x, y)$ con el plano $y = y_0$. Esta es una curva que en el punto $x = x_0$ tiene una recta tangente cuya pendiente es precisamente $\varphi'(x_0) = \frac{\partial f}{\partial x}(\mathbf{p})$

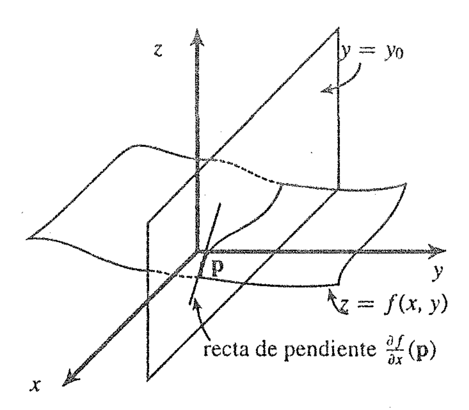


Figura 1. La derivada parcial de f con respecto a x.

Análogamente, la función $\psi(y) = f(x_0, y)$ representa la curva obtenida de la intersección de la superficie $z = f(x, y)$ con el plano $x = x_0$, cuya pendiente en el punto $y = y_0$ es $\psi'(y_0) = \frac{\partial f}{\partial y}(\mathbf{p})$ (figura 2).

De esta manera, las derivadas parciales de una función $z = f(x, y)$ en un punto $\mathbf{p} = (x_0, y_0)$ nos hablan del comportamiento geométrico (la inclinación) de la superficie que tal función representa, en *las direcciones de los ejes* x e y. Esta es pues una información "parcial".

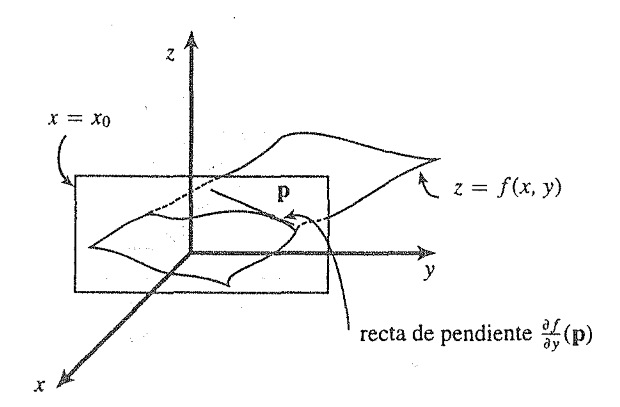


Figura 2. La derivada parcial de f con respecto a y.

NOTA: Las derivadas parciales son conceptos "puntuales", es decir, se habla de la derivada parcial de una función en *un punto dado* (de su dominio). Es por eso que, en general, se debe hacer explícito el punto $\mathbf{p} = (x_0, y_0)$ donde están evaluados $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$, escribiendo $\frac{\partial f}{\partial x}(x_0, y_0)$ o $\frac{\partial f}{\partial x}(\mathbf{p})$ (igualmente para $\frac{\partial f}{\partial y}$). Sin embargo, muchas veces calculamos las derivadas parciales "en un punto cualquiera (x, y) de su dominio". En tal caso, basta escribir $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ (o f_x y f_y, o f_1 y f_2, o $D_1 f$ y $D_2 f$, respectivamente).

Ejemplo 1. Sea $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 y^3$. Entonces

$$\begin{aligned}\frac{\partial f}{\partial x} &= \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h} = \lim_{h\to 0} \frac{(x+h)^2 y^3 - x^2 y^3}{h} \\ &= \lim_{h\to 0}(2xy^3 + hy^3) = 2xy^3\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial y} &= \lim_{h\to 0} \frac{f(x, y+h) - f(x, y)}{h} = \lim_{h\to 0} \frac{x^2(y+h)^3 - x^2 y^3}{h} \\ &= \lim_{h\to 0}(3x^2y^2 + 3x^2yh + x^2h^3) = 3x^2y^2\end{aligned}$$

En este ejemplo se observa que, como era de esperarse, las derivadas parciales de una función $z = f(x, y)$ se obtienen "derivando parcialmente" cada una de las variables, y conservando la otra como constante (es decir, pensando en la función f como dependiente sólo de x o de y) ■

Ejemplo 2. Sea $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \operatorname{sen}\left(\sqrt{2x^3 + y^2}\right)$. Entonces

$$\frac{\partial f}{\partial x} = \left(\cos\sqrt{2x^3 + y^2}\right)\frac{1}{2\sqrt{2x^3 + y^2}}(6x^2) = \frac{3x^2 \cos\sqrt{2x^3 + y^2}}{\sqrt{2x^3 + y^2}}$$

$$\frac{\partial f}{\partial y} = \left(\cos\sqrt{2x^3 + y^2}\right)\frac{1}{2\sqrt{2x^3 + y^2}}(2y) = \frac{y\cos\sqrt{2x^3 + y^2}}{\sqrt{2x^3 + y^2}}$$ ■

Ejemplo 3. Sea $f(x, y) = x^2 + y^2$. Las derivadas parciales de esta función en el punto $(1, 0)$ son

$$\frac{\partial f}{\partial x}(1, 0) = 2x\Big|_{(1,0)} = 2(1) = 2 \qquad \frac{\partial f}{\partial y}(1, 0) = 2y\Big|_{(1,0)} = 2(0) = 0$$

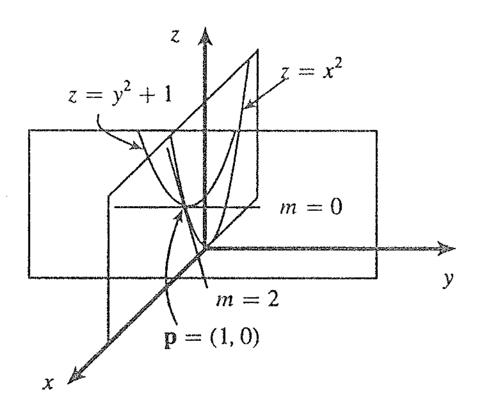


Figura 3. Las derivadas parciales de la función $z = x^2 + y^2$ en p. ■

Esto significa que la curva $f(x, 0) = x^2$ (que está en el plano $y = 0$) tiene pendiente $m = 2$ en $x = 1$, y que la curva $f(1, y) = 1 + y^2$ (que está en el plano $x = 1$) tiene pendiente $m = 0$ en $y = 0$ (figura 3).

Al considerar el caso general de funciones de n variables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, conviene usar notación vectorial para definir las derivadas parciales. El esquema que sigue la definición de derivada parcial de f respecto de su i-ésima variable, digamos x_i, es el mismo: digamos que se toma la función f con su i-ésima variable incrementada en h. Se resta la función evaluada en el punto considerado, se divide entre el valor del incremento (h) y se toma el límite cuando éste tiende a cero.

Sean $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ los vectores de la base canónica de $\mathbb{R}^n$. Entonces $h\mathbf{e}_i$ es el vector que tiene sólo ceros, excepto en la i-ésima coordenada, donde tiene h. Si $\mathbf{x} \in \mathbb{R}^n$ es un vector cualquiera, el vector $\mathbf{x} + h\mathbf{e}_i$ será idéntico a $\mathbf{x}$, excepto en la i-ésima coordenada, donde aparece la i-ésima coordenada de $\mathbf{x}$, x_i, incrementada en h.

Definición. (derivadas parciales). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$ y sea $\mathbf{x}_0 \in U$.[1] Se define la derivada parcial de f con respecto a su i-ésima variable en el punto $\mathbf{x}_0$, denotada por $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$ (o $f_{x_i}(\mathbf{p})$, o $f_i(\mathbf{p})$, o $D_i f(\mathbf{p})$), como el límite

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{e}_i) - f(\mathbf{x}_0)}{h}$$

cuando éste existe. ■

Así, por ejemplo, para la función $f(x, y, z)$ tenemos

$$\frac{\partial f}{\partial z} = \lim_{h \to 0} \frac{f(x, y, z + h) - f(x, y, z)}{h}$$

y para la función $g(x, y, z, u)$ tenemos

$$\frac{\partial f}{\partial y} = \lim_{h \to 0} \frac{g(x, y + h, z, u) - g(x, y, z, u)}{h}$$

etc.

[1] La razón de definir las derivadas parciales en un punto perteneciente a un conjunto abierto U (dominio de la función), es para poder asegurar que para $h \in \mathbb{R}$ pequeño, se tenga $x_0 + h\mathbf{e}_i \in U$ y que así haga sentido la expresión $f(x_0 + h\mathbf{e}_i)$, la cual aparece en la definición de derivada parcial.

Es claro que también en este caso general las derivadas parciales de una función, respecto a alguna de sus variables, se calculan derivando la función pensando ésta como si dependiera sólo de la variable en cuestión, con las variables restantes fijas.

Ejemplo 4. Sea $f\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = e^{x^3y^4z^5}$. Entonces

$$\frac{\partial f}{\partial x} = e^{x^3y^4z^5}(3x^2y^4z^5), \quad \frac{\partial f}{\partial y} = e^{x^3y^4z^5}(4x^3y^3z^5), \quad \frac{\partial f}{\partial z} = e^{x^3y^4z^5}(5x^3y^4z^4)$$ ■

Recordando que para una función $y = \varphi(x)$ de una variable real, la derivada $\varphi'(x_0)$ mide la velocidad de variación de y con respecto a x en $x = x_0$, tenemos, de manera análoga, que para la función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, su derivada parcial respecto de su i-ésima variable en $\mathbf{p} \in U$, $\frac{\partial f}{\partial x_i}(\mathbf{p})$, mide la velocidad de variación *parcial* de la función con respecto de dicha variable, cuando las demás se mantienen fijas.

Ejemplo 5. El volumen V de un gas encerrado en un recipiente elástico es función de su presión P y de su temperatura T, según la ley $V = \frac{KT}{P}$, en donde K es cierta constante. Entonces

$$\frac{\partial V}{\partial T} = \frac{K}{P} \quad \text{y} \quad \frac{\partial V}{\partial P} = -\frac{KT}{P^2}$$

miden las variaciones del volumen con la temperatura (a presión constante) y del volumen con la presión (a temperatura constante), de modo respectivo. ■

Retomando el interés de llegar a establecer una noción de diferenciabilidad equivalente a la de (simplísima) para funciones de una variable (en cuyo caso diferenciabilidad equivale a existencia de la derivada), nos podríamos preguntar si la sola existencia de las derivadas parciales de una función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto $\mathbf{x}_0 \in U$ nos puede dar un concepto de diferenciabilidad como el requerido. Es decir, si establecemos la definición: "la función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$ es diferenciable en el punto $\mathbf{x}_0 \in U$ si las derivadas parciales $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$, $i = 1, 2, \ldots, n$ existen", nos preguntamos si esta noción de diferenciabilidad es equivalente a la de funciones de una sola variable.

No podemos negar que esta "definición" de diferenciabilidad para funciones de varias variables goza de una gran simplicidad. Sería bueno que esta fuera en verdad la definición que buscamos. Sin embargo, este anhelo pronto se viene abajo pues es fácil convencerse que la existencia de las derivadas parciales de una función, (siendo una condición necesaria) está muy lejos de ser una condición suficiente para que la función sea diferenciable en el sentido buscado.

Para poder ver esto, recordemos que, en el caso de funciones de una variable, la diferenciabilidad de una función en un punto implica la continuidad de la función en ese punto. De modo que, en el caso de varias variables, la noción de diferenciabilidad (que procuramos sea equivalente al caso de una variable, v.gr. que tenga las mismas propiedades) que buscamos, debe respetar esta propiedad (i.e. diferenciabilidad debe implicar continuidad). El ejemplo siguiente nos desengaña sobre la posible validez de la definición establecida anteriormente.

Ejemplo 6. Considere la función $f\colon \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

Calculemos $\frac{\partial f}{\partial x}(0,0)$ y $\frac{\partial f}{\partial y}(0,0)$. Según la definición de derivadas parciales, tenemos

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{h(0)}{h^2+0^2} - 0}{h} = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0} \frac{f(0,h)-f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{(0)h}{h^2+0^2} - 0}{h} = 0$$

de modo que la función tiene ambas derivadas parciales (iguales a cero) en el origen. Sin embargo, en el ejemplo 9 de la sección anterior, vimos que esta función es discontinua en (0, 0). ■

En la próxima sección estudiaremos un concepto más fuerte que el de derivada parcial de una función de varias variables, con el que trataremos de reformular la definición fallida de diferenciabilidad establecida en esta sección.

Ejercicios (Capítulo 2, Sección 4)

En los ejercicios 1–7, identifique las expresiones dadas como derivadas parciales de funciones de varias variables respecto alguna de sus variables. Obtenga la derivada parcial indicada.

1. $\lim_{h\to 0} \frac{(x+h)^4 y^5 - x^4 y^5}{h}$

2. $\lim_{h\to 0} \frac{3y^2 \operatorname{sen}(x+h)^2 + \tan^2(x+h) - 3y^2 \operatorname{sen} x^2 - \tan^2 x}{h}$

3. $\lim_{h\to 0} \frac{\ln \frac{y+h}{x} + 3\ln \frac{x}{y+h} - \ln \frac{y}{x} - 3\ln \frac{x}{y}}{h}$

4. $\lim_{h\to 0} \frac{\sqrt{(x+h)y \operatorname{sen} z} - \sqrt{xy \operatorname{sen} z}}{h}$

5. $\lim_{h\to 0} \frac{\left(\frac{x}{z+h} + \frac{z+h}{y} + (z+h)\cos^5(z+h)^4\right)^{1/3} - \left(\frac{xy+z^2}{zy} + z\cos^5 z^4\right)^{1/3}}{h}$

6. $\lim_{h\to 0} \frac{e^{x^2y^2z^2} e^{2xhy^2z^2} (e^{hyz})^2 - e^{(xyz)^2}}{h}$

7. $\lim_{h\to 0} \frac{y \operatorname{sen} xz(\cos xh - 1) + y\cos xz \operatorname{sen} xh}{h}$

En los ejercicios 8–30 obtenga todas las derivadas parciales de las funciones indicadas.

8. $f(x, y) = (4x^2y^4 - 3x^2 + 8y^3)^3$

9. $f(x, y) = \operatorname{arcsen} \frac{y}{x} + \arccos \frac{x}{y}$

10. $f(x, y) = \frac{x+y}{x-y}$

11. $f(x, y) = x + y + xy + \frac{x}{y} + \frac{y}{x}$

12. $f(x, y) = x^y + y^x$

13. $f(x, y) = x^x + y^y + x^y y^x$

14. $f(x, y) = (2x + 3y)^x + (2x + 3y)^y$

15. $f(x, y) = x^{y^x} + y^{x^y} + (x^y)^x (y^x)^y$

16. $f(x, y) = \dfrac{1}{\ln^2(1 + x^2 + y^2)}$

17. $f(x, y) = (2y)^x + 2^y$

18. $f(x, y) = x \ln y - y \ln x$

19. $f(x, y) = \arccos \sqrt{x^2 + y^2}$

20. $f(x, y) = \arctan(2x + 3 \operatorname{sen} x)^y$

21. $f(x, y) = \ln \dfrac{1 - (x^2 + y^2)^{1/2}}{1 + (x^2 + y^2)^{1/2}}$

22. $f(x, y, z) = \ln \dfrac{1 - (x^2 + y^2 + z^2)^{1/2}}{1 + (x^2 + y^2 + z^2)^{1/2}}$

23. $f(x, y, z) = x^y + x^z + y^x + y^z + z^x + z^y$

24. $f(x, y, z) = (x^y)^z + (x^z)^y + (y^x)^z + (y^z)^x + (z^x)^y + (z^y)^x$

25. $f(x, y, z) = x^2 \arctan \sqrt{1 + y + z} + \ln(1 + y + z)$

26. $f(x, y, z) = xy \operatorname{sen} z + xz \cos y + yz \tan x$

27. $f(x, y, z) = x^2 y^3 z^4 \operatorname{sen}^2 x \cos^3 y \tan^4 z$

28. $f(x, y, z) = \ln\left(\dfrac{x}{y} + \dfrac{x}{z} + \dfrac{y}{x} + \dfrac{y}{z} + \dfrac{z}{x} + \dfrac{z}{y}\right)$

29. $f(x, y, z, u) = x^{y+z+u} z^{x+y+u}$

30. $f(x_1, x_2, \ldots, x_n) = \dfrac{x_1}{x_2} + \dfrac{x_2}{x_3} + \cdots + \dfrac{x_{n-1}}{x_n}$

31. Sea $f(x, y, z) = x^{y^z} + x^{z^y} + y^{x^z} + y^{z^x} + z^{x^y} + z^{y^x}$. Calcule las derivadas parciales de esta función en el punto $(1, 1, 1)$.

32. Sea $f(x, y, z) = x^{(y/z)} + x^{(z/y)} + y^{(x/z)} + y^{(z/x)} + z^{(x/y)} + z^{(y/x)}$. Calcule las derivadas parciales de esta función en el punto $(1, 1, 1)$.

33. Sea $f(x_1, x_2, \ldots, x_n) = \ln(x_1 x_2 \ldots x_n)$. Calcule

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(1, 1, \ldots, 1)$$

34. Sea $f(x, y) = 3x^2 y^4 - 12x^6 + 2xy^5$. Verifique que

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 6f(x, y)$$

35. Considere la función del ejercicio 30

$$f(x_1, x_2, \ldots, x_n) = \frac{x_1}{x_2} + \frac{x_2}{x_3} + \cdots + \frac{x_{n-1}}{x_n}$$

Demuestre que $\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = 0$.

36. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$. Demuestre que si la derivada parcial de f respecto de x existe en el punto $\mathbf{p} = (x_0, y_0) \in U$, entonces f es continua respecto de su primera variable en $\mathbf{p}$ (ver ejercicio 80 de la sección anterior). Establezca un resultado análogo para la segunda variable de la función f. Use la función $f(x, y) = \sqrt{x^2 + y^2}$ (en el punto $\mathbf{p} = (0, 0)$) para probar que las afirmaciones recíprocas son falsas.

37. Sean $f, g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ dos funciones definidas en el conjunto abierto U de $\mathbb{R}^2$ que tienen derivadas parciales en U.

a. Demuestre que la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F = f + g$ tiene derivadas parciales en U y que éstas son

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \qquad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} + \frac{\partial g}{\partial y}$$

b. Demuestre que la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F = fg$ tiene derivadas parciales en U y que éstas son

$$\frac{\partial F}{\partial x} = f\frac{\partial g}{\partial x} + g\frac{\partial f}{\partial x} \qquad \frac{\partial F}{\partial y} = f\frac{\partial g}{\partial y} + g\frac{\partial f}{\partial y}$$

c. Demuestre que si $g(x, y) \neq 0$ en U, la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F = \frac{f}{g}$ tiene derivadas parciales en U y que éstas son

$$\frac{\partial F}{\partial x} = \frac{g\frac{\partial f}{\partial x} - f\frac{\partial g}{\partial x}}{g^2} \qquad \frac{\partial F}{\partial y} = \frac{g\frac{\partial f}{\partial y} - f\frac{\partial g}{\partial y}}{g^2}$$

38. a. Calcule la pendiente de la recta tangente a la curva de intersección de la superficie $z = x^3y + 5y^2$ con el plano $x = 2$, en el punto en el que $y = 1$.

b. Calcule la pendiente de la recta tangente a la curva de intersección de la superficie $z = x^2 + y^3x$ con el plano $y = 2$, en el punto en el que $x = 2$.

39. Demuestre que la ecuación de la recta tangente (cuando existe) a la curva de intersección de la gráfica de la función $z = f(x, y)$ con el plano $y = y_0$ en el punto en el que $x = x_0$ viene dada por

$$\begin{cases} x = x_0 + t \\ y = y_0 \\ z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)t \end{cases} \qquad t \in \mathbb{R}$$

40. Demuestre que la ecuación de la recta tangente (cuando existe) a la curva de intersección de la gráfica de la función $z = f(x, y)$ con el plano $x = x_0$ en el punto en el que $y = y_0$ viene dada

por

$$\begin{cases} x = x_0 \\ y = y_0 + t \\ z = f(x_0, y_0) + \dfrac{\partial f}{\partial x}(x_0, y_0)t \end{cases} \qquad t \in \mathbb{R}$$

41. Sea $g: \mathbb{R} \to \mathbb{R}$ una función continua y positiva definida en $\mathbb{R}$. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = \int_x^y g(t)\, dt$$

a. ¿Para qué puntos $(x, y) \in \mathbb{R}^2$ se tiene que $f(x, y) > 0$?

b. ¿Para qué puntos $(x, y) \in \mathbb{R}^2$ se tiene que $f(x, y) < 0$?

c. ¿Cuál es el nivel cero de $f(x, y)$?

d. Calcule las derivadas parciales de la función f.

42. Resuelva el ejercicio anterior suponiendo que la función g es impar y que $g(t) > 0$ para $t > 0$.

Calcule las derivadas parciales de cada una de las funciones de los ejercicios 43–50, donde $g: \mathbb{R} \to \mathbb{R}$ una función continua.

43. $f(x, y) = \int_x^{xy} g(t)\, dt$

44. $f(x, y) = \int_{x+y}^{x-y} g(t)\, dt$

45. $f(x, y) = \int_{xy}^{y} (x^2 + y^2) g(t)\, dt$

46. $f(x, y) = \int_{x^y}^{y^x} g(t)\, dt$

47. $f(x, y) = \int_{\int_1^x g(t)\, dt}^{y} g(t)\, dt$

48. $f(x, y) = \int_{\int_y^x g(t)\, dt}^{\int_x^y g(t)\, dt} g(t)\, dt$

49. $f(x, y, z) = \int_{xyz}^{x+y+z} g(t)\, dt$

50. $f(x, y, z) = \int_{x+y+z}^{\int_{x+y+z}^{\int_x^y g(t)\, dt} g(t)\, dt} g(t)\, dt$

Para cada una de las funciones dadas en los ejercicios 51—71, en las que $g, h: \mathbb{R} \to \mathbb{R}$ son funciones definidas en $\mathbb{R}$, diferenciables (es decir, tal que $g'(t)$ y $h'(t)$ existen para todo $t \in \mathbb{R}$), calcule sus derivadas parciales.

51. $f(x, y) = g(x) + 5h(y)$

52. $f(x, y) = 2g(x)h(y) + g^2(x) + h(y^2)$

53. $f(x, y) = \dfrac{1 + h(x)}{1 + (g(y))^2}$

54. $f(x, y) = h(x)g(h(y)) + g(y)h(g(x))$

55. $f(x, y) = g(h(x)) \operatorname{sen} h(g(y)) + h(g(y)) \cos g(h(x))$

56. $f(x, y, z) = \ln(1 + g^2(x) + h^4(y) + g^6(z))$

57. $f(x, y, z) = g(z)h(g(x)h(y))$

58. $f(x, y, z) = g(g(x)g(g(y)g(h(z))))$

59. $f(x, y, z) = (g(x))^{h(y)} + (h(y))^{g(z)}$

60. $f(x, y, z) = (g(x))^{(h(y))^{g(z)}}$

61. $f(x, y, z) = (g(h(x)))^{(h(g(y)))^{g(h(z))}}$

62. $f(x, y) = (1 + g^2(x))^{h(x)}(1 + h^4(y))^{h(y)}$

63. $f(x, y) = (1 + g^2(x) + h^2(y))^{g^3(x)}$

64. $f(x, y) = (g^2(h^3(y)))^{h^3(g^2(y))}$

65. $f(x, y) = (1 + g^2(x))^{(1+h^2(y))^{g(h(x))}}$

66. $f(x, y) = (\ln(1 + x^2))^{(\ln(1+g^2(x))^{h^2(y)}}$

67. $f(x, y, z) = yz \operatorname{sen}(1 + h^2(x))^{(x^2+1)}$

68. $f(x, y, z) = xyz(1 + z^2)^{g(x)h(z)}$

69. $f(x, y, z) = \left(g^2(h(x))h^2(g(z))\right)^{h^2(g(y))g^2(h(z))}$

70. $f(x, y, z) = \left(\ln(1 + g^2(x) + h^4(y) + g^6(z))\right)^{g^2(x)h^4(y)g^6(z)}$

71. $f(x, y, z) = (1 + g^2(x)h^4(y)g^6(z))^{\ln(1+g^2(x)h^4(y)g^6(z))}$

72. Sea $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$, con rango $J \subseteq \mathbb{R}$, que tiene derivadas parciales en el punto $\mathbf{p} = (x_0, y_0)$ de U. Sea $g\colon I \subseteq \mathbb{R} \to \mathbb{R}$ una función real definida en el intervalo abierto I que contiene a J, diferenciable en el punto $g(f(\mathbf{p}))$. Queremos ver que la función compuesta $g \circ f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $(g \circ f)(x, y) = g(f(x, y))$ tiene derivadas parciales en el punto $\mathbf{p}$, y queremos obtener fórmulas para calcularlas. Los siguientes pasos son el "guión" de un argumento que *hace plausible* el resultado que se quiere obtener. Se pide al lector que justifique cada paso.

$$
\begin{aligned}
\frac{\partial}{\partial x}(g \circ f)(x_0, y_0) &= \lim_{h \to 0} \frac{(g \circ f)(x_0 + h, y_0) - (g \circ f)(x_0, y_0)}{h} \\
&= \lim_{h \to 0} \frac{g(f(x_0 + h, y_0)) - g(f(x_0, y_0))}{h} \\
&\xrightarrow{\text{Sea } k = f(x_0+h, y_0) - f(x_0, y_0)} \\
&= \lim_{h \to 0} \frac{g(f(x_0, y_0) + k) - g(f(x_0, y_0))}{h} \\
&= \lim_{h \to 0} \frac{g(f(x_0, y_0) + k) - g(f(x_0, y_0))}{k} \frac{k}{h} \\
&= \lim_{k \to 0} \frac{g(f(x_0, y_0) + k) - g(f(x_0, y_0))}{k} \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \\
&= g'(f(x_0, y_0)) \frac{\partial f}{\partial x}(x_0, y_0)
\end{aligned}
$$

Un resultado análogo se obtiene para la derivada parcial de $g \circ f$ respecto de y. En resumen, se tienen las fórmulas

$$\frac{\partial}{\partial x}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial f}{\partial x}(\mathbf{p})$$
$$\frac{\partial}{\partial y}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial f}{\partial y}(\mathbf{p})$$

Así, una derivada parcial de la función compuesta $g \circ f$, de la función de una sola variable $g(t)$ con la función de dos variables $f(x, y)$, se obtiene multiplicando la derivada de g (evaluada en $f(x, y)$) y multiplicándola por la derivada parcial correspondiente de la función f. Por ejemplo, para cualquier g diferenciable definida en $\mathbb{R}$ se tiene

$$\frac{\partial}{\partial x}g(x^2 + 4xy^3) = (2x + 4y^3)g'(x^2 + 4y^3)$$
$$\frac{\partial}{\partial y}g(x^2 + 4xy^3) = 12xy^2g'(x^2 + 4y^3)$$

Siga paso a paso el argumento presentado en este ejercicio con el que se "demuestran" las fórmulas de las derivadas parciales de la composición y diga dónde falla el argumento (sugerencia: siga paso a paso el argumento con la función $f(x, y) = y$). En el capítulo siguiente se verá (bien demostrado) un resultado más general que el que aquí se presenta. La finalidad de este ejercicio es que desde este momento comencemos a hacer uso de este caso sencillo de derivación de funciones compuestas.

Sea g una función real diferenciable de una sola variable real. Para cada una de las funciones dadas en los ejercicios 73–80, determine sus derivadas parciales.

73. $F(x, y) = g(xy)$

74. $F(x, y) = g(3x^2 + 7y^2)$

75. $F(x, y) = g^2(x + y)$

76. $F(x, y) = g(x + y^2) + g(x^2 + y)$

77. $F(x, y) = g(3x^3y^4)g(3x^3 + y^4)$

78. $F(x, y) = \ln(4 + g^4(ax + by + c))$

79. $F(x, y) = \arctan(1 + \operatorname{sen} g(x \cos y + y \operatorname{sen} x))$

80. $F(x, y) = g^7(g^6(x^5 + y^4)g^3(x^2 + y))$

En los ejercicios 81–85, $z = \phi(x, y)$ es una función real de variable real, diferenciable en $\mathbb{R}$. Demuestre que la función dada satisface la expresión indicada.

81. $f(x, y) = x^2\phi(x^2y)$, $\quad x\dfrac{\partial f}{\partial x} - 2y\dfrac{\partial f}{\partial y} = 2z$

82. $f(x, y) = y\phi(x + y)$, $\quad y\left(\dfrac{\partial f}{\partial y} - \dfrac{\partial f}{\partial x}\right) = z$

83. $f(x, y) = x^2\phi(3x + y^2)$, $\quad 2xy\dfrac{\partial f}{\partial x} - 3x\dfrac{\partial f}{\partial y} = 4yz$

84. $f(x, y) = x\phi(xy^3)$, $\quad 3x\dfrac{\partial f}{\partial x} - y\dfrac{\partial f}{\partial y} = 3z$

85. $f(x, y) = e^{x+y}\phi(xe^y)$, $\quad x\dfrac{\partial f}{\partial x} - \dfrac{\partial f}{\partial y} = z(x-1)$

2.5 Derivadas direccionales

Consideremos la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $\mathbf{v} \in \mathbb{R}^n$ un vector dado de $\mathbb{R}^n$, cuya norma es 1. Ahora queremos estudiar la variación de la función f en el punto $\mathbf{x}_0 \in U$ cuando su argumento varía en la dirección marcada por el vector $\mathbf{v}$. La idea para lograr esto será la misma que aparece en el concepto de derivada de una función de una variable, y, más recientemente, en el estudio de las derivadas parciales en la sección anterior, a saber, la derivada será "el límite cuando el incremento de la variable tiende a cero del cociente del incremento de la función, dividido entre el incremento de la variable". En el caso de las derivadas parciales, "el incremento de la variable" correspondía al de la variable respecto de la cual se estaba derivando. Lo que haremos ahora será tomar el "incremento de la variable", comenzando en el punto $\mathbf{x}_0 \in U$, y yendo en la dirección del vector unitario $\mathbf{v} \in \mathbb{R}^n$ dado. Esquemáticamente

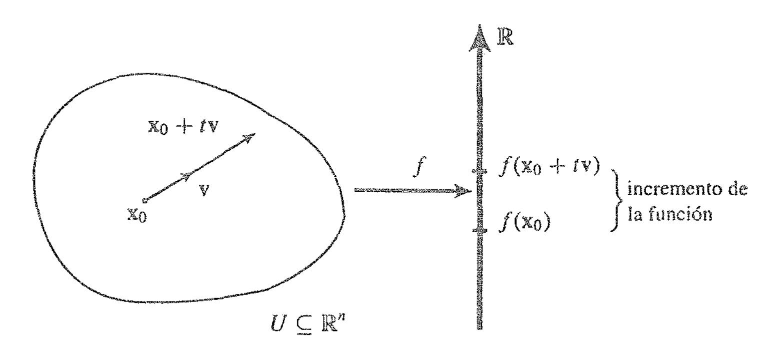


Figura 1. Variación de la función en la dirección de $\mathbf{v}$.

Nótese que (en el esquema anterior) en este caso la variable (vectorial) $\mathbf{x} \in U$ se mueve t unidades en dirección del vector $\mathbf{v}$, pasando del punto $\mathbf{x}_0 \in U$ al punto $\mathbf{x}_0 + t\mathbf{v}$ (de otro modo, la magnitud de esta variación es $\|t\mathbf{v}\| = |t|\|\mathbf{v}\| = |t|$). Así, tomando el cociente del incremento de la función $f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)$, dividiendo entre t y tomando el límite cuando t tiende a cero, obtenemos la manera de variar de f en $\mathbf{x}_0$ en la dirección de $\mathbf{v}$. A esto lo llamamos derivada direccional de f en $\mathbf{x}_0$ en la dirección de $\mathbf{v}$. Incluso se tiene

Definición. (derivada direccional) Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$ y sea $\mathbf{x}_0 \in U$ un punto dado de U. Sea $\mathbf{v} \in \mathbb{R}^n$ un vector unitario dado. Se define la derivada de la función f en $\mathbf{x}_0$, en la dirección del vector $\mathbf{v}$, denotada por $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$, o $D_{\mathbf{v}}f(\mathbf{x}_0)$, como el límite

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$ ■

La primera observación que debemos hacer de la definición anterior es que el concepto de derivada direccional es un concepto que generaliza el de derivada parcial estudiado en la sección 2.4. En efecto, si $\mathbf{v} = \mathbf{e}_i =$ i-ésimo vector de la base canónica de $\mathbb{R}^n$, tenemos que, efectivamente $\|\mathbf{v}\| = \|\mathbf{e}_i\| = 1$, y

$$\frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{x}_0) = \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_i) - f(\mathbf{x}_0)}{t} = \frac{\partial f}{\partial x_i}(\mathbf{x}_0)$$
$$= \text{derivada parcial de } f \text{ en } \mathbf{x}_0 \text{ respecto de su } i\text{-ésima variable}$$

la cual, dijimos ya, medía la variación de la función f en $\mathbf{x}_0$ en la dirección del vector $\mathbf{e}_i$.

Otro punto que debemos observar es lo que pasa en el (muy frecuente —en este libro) caso de una función de dos variables $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$. Podemos escribir en general el vector unitario $\mathbf{v} \in \mathbb{R}^2$ como

$$\mathbf{v} = (\cos\theta, \operatorname{sen}\theta),\ 0 \le \theta \le 2\pi$$

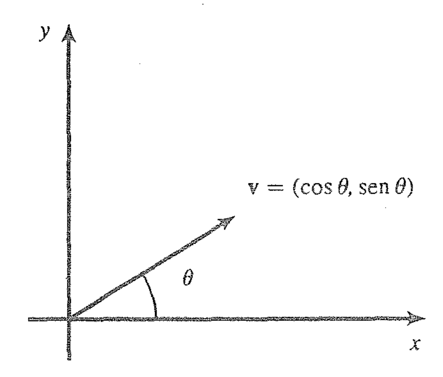


Figura 2. Un vector unitario en $\mathbb{R}^2$.

de modo que la derivada direccional de f en el punto $(x_0, y_0) \in U$ en la dirección de $\mathbf{v}$ se vería como

$$\frac{\partial f}{\partial \mathbf{v}}(x_0, y_0) = \lim_{t\to 0} \frac{f((x_0, y_0) + t(\cos\theta, \operatorname{sen}\theta)) - f(x_0, y_0)}{t}$$
$$= \lim_{t\to 0} \frac{f(x_0 + t\cos\theta, y_0 + t\operatorname{sen}\theta) - f(x_0, y_0)}{t}$$

Nuevamente observamos que si $\theta = 0$, o bien si $\mathbf{v} = \mathbf{i} = (1, 0)$, entonces

$$\frac{\partial f}{\partial \mathbf{v}}(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + t\cos 0, y_0 + t\operatorname{sen} 0) - f(x_0, y_0)}{t}$$
$$= \lim_{t\to 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} = \frac{\partial f}{\partial x}(x_0, y_0)$$

Del mismo modo, si $\theta = \pi/2$, o bien, si $\mathbf{v} = \mathbf{j} = (0, 1)$, entonces

$$\frac{\partial f}{\partial \mathbf{v}}(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + t\cos\frac{\pi}{2}, y_0 + t\operatorname{sen}\frac{\pi}{2}) - f(x_0, y_0)}{t}$$
$$= \lim_{t\to 0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t} = \frac{\partial f}{\partial y}(x_0, y_0)$$

Veamos ahora algunos ejemplos.

Ejemplo 1. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ la función $f(x, y) = x^2 + y^2$. Calculemos la derivada de esta función en un punto $(x, y) \in \mathbb{R}^2$ en la dirección del vector $\mathbf{v} = (\cos\theta, \operatorname{sen}\theta)$. Según la discusión anterior tenemos

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}} &= \lim_{t\to 0} \frac{f(x + t\cos\theta, y + t\operatorname{sen}\theta) - f(x, y)}{t}\\ &= \lim_{t\to 0} \frac{(x + t\cos\theta)^2 + (y + t\operatorname{sen}\theta)^2 - (x^2 + y^2)}{t}\\ &= \lim_{t\to 0}(2x\cos\theta + 2y\operatorname{sen}\theta + t) = 2x\cos\theta + 2y\operatorname{sen}\theta\end{aligned}$$

También, como dijimos, si $\theta = 0$ se tiene

$$\frac{\partial f}{\partial \mathbf{v}} = 2x = \frac{\partial f}{\partial x}$$

y si $\theta = \pi/2$

$$\frac{\partial f}{\partial \mathbf{v}} = 2y = \frac{\partial f}{\partial y}$$

■

Ejemplo 2. Sea $f: \mathbb{R}^2 \to \mathbb{R}$, la función definida como

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0)\\ 0 & \text{si } (x, y) = (0, 0)\end{cases}$$

(ver ejemplo 6 de la sección 2.4). Calculemos $\dfrac{\partial f}{\partial \mathbf{v}}(0, 0)$

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(0, 0) &= \lim_{t\to 0} \frac{f(t\cos\theta, t\operatorname{sen}\theta) - f(0, 0)}{t}\\ &= \lim_{t\to 0} \frac{\dfrac{(t\cos\theta)(t\operatorname{sen}\theta)}{(t\cos\theta)^2 + (t\operatorname{sen}\theta)^2}}{t} = \lim_{t\to 0} \frac{\cos\theta\operatorname{sen}\theta}{t}\end{aligned}$$

Se ve entonces que este límite existe si y sólo si $\cos\theta\operatorname{sen}\theta = 0$, i.e. si $\theta = 0, \pi/2, \pi, 3\pi/2$. En tal caso $\frac{\partial f}{\partial \mathbf{v}}(0, 0) = 0$. Nótese entonces que esta función *no tiene* derivadas direccionales (en el origen) excepto en las "direcciones canónicas" (las de los ejes x y y). Es decir, esta función tiene derivadas parciales iguales a cero en $(0, 0)$, pero no tiene derivadas en ninguna otra dirección. ■

Ejemplo 3. Sea $f: \mathbb{R}^3 \to \mathbb{R}$, la función $f(x, y, z) = 2x^3 + 7y^2 + 9z^2$ y sea $\mathbf{v} = (a, b, c)$ un vector unitario dado en el espacio $\mathbb{R}^3$. La derivada direccional de esta función en un punto $(x, y, z) \in \mathbb{R}^3$ arbitrario en la dirección de $\mathbf{v}$ es

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}} &= \lim_{t\to 0} \frac{f((x, y, z) + t(a, b, c)) - f(x, y, z)}{t}\\ &= \lim_{t\to 0} \frac{f(x + ta, y + tb, z + tc) - f(x, y, z)}{t}\end{aligned}$$

$$
\begin{aligned}
&= \lim_{t\to 0} \frac{2(x+ta)^3 + 7(y+tb)^2 + 9(z+tc)^2 - (2x^3 + 7y^2 + 9z^2)}{t} \\
&= \lim_{t\to 0}(6x^2a + 6xta^2 + 2t^2a^3 + 14yb + 7tb^2 + 18zc + 9zc^2) \\
&= 6x^2a + 14yb + 18zc
\end{aligned}
$$
■

Podemos interpretar geométricamete la derivada direccional de una función de dos variables $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ en un punto $(x_0, y_0) \in U$. Sin pérdida de generalidad supongamos que $(x_0, y_0) = (0, 0)$. Sea $\mathbf{v} = (\cos\theta, \operatorname{sen}\theta) \in \mathbb{R}^2$ el vector unitario en la dirección del cual calculamos la derivada de la función f en el origen. Consideremos el plano

$$x \operatorname{sen}\theta - y\cos\theta = 0$$

Este es un plano perpendicular al plano $z = 0$ que contiene al vector $\mathbf{v}$. La intersección de este plano con la superficie $z = f(x, y)$ es una curva en el espacio. La derivada direccional de f en $(0, 0)$ en la dirección de $\mathbf{v}$ es la pendiente de la recta tangente a esta curva en $(0, 0)$.

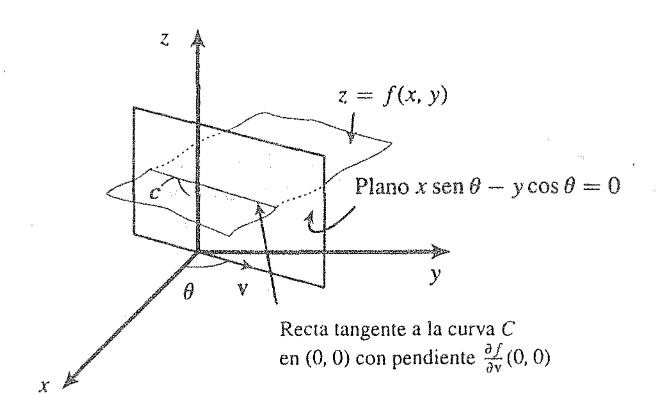


Figura 3. Interpretación geométrica de la derivada direccional.

NOTA: El concepto de pendiente de una recta está totalmente comprometido con "las unidades que se siembran" en los ejes coordenados. Al decir que la derivada direccional $\frac{\partial f}{\partial \mathbf{v}}(0, 0)$ es la pendiente de la recta tangente a la curva de intersección de la superficie $z = f(x, y)$ con el plano $x \operatorname{sen}\theta - y\cos\theta = 0$ (que contiene al vector $\mathbf{v}$) en $(0, 0)$, estamos pensando en la pendiente medida en el plano zs, en que s es el eje generado por el vector $\mathbf{v}$ cuya magnitud marca la unidad que se toma en ese eje (ver la figura 4). El siguiente ejemplo muestra en un caso concreto esta observación.

Ejemplo 4. Considere la función $f\colon \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = (x-1)^2 + (y-1)^2$. La superficie que esta función representa es la misma que la del paraboloide $z = x^2 + y^2$, recorrido su vértice al punto $(1, 1, 0)$. Calculemos la derivada direccional de esta función en $(0, 0)$ en la dirección del

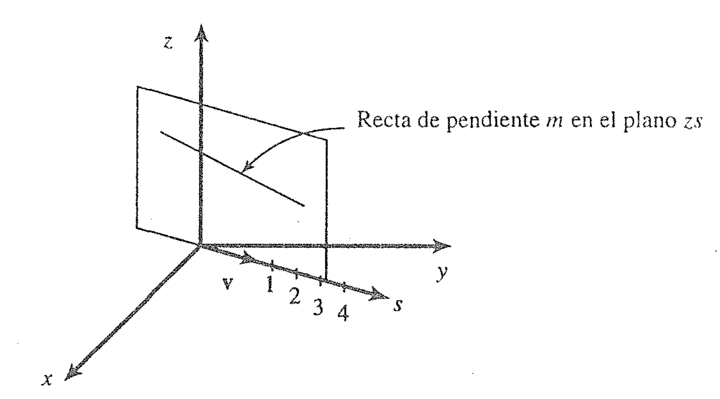


Figura 4. Las unidades utilizadas para medir la pendiente se toman sobre el plano zs, donde la magnitud de v da la unidad.

vector unitario $\mathbf{v} = (\sqrt{2}/2, \sqrt{2}/2)$. Tenemos

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(0,0) &= \lim_{t\to 0} \frac{f(\sqrt{2}t/2, \sqrt{2}t/2) - f(0,0)}{t}\\ &= \lim_{t\to 0} \frac{(\sqrt{2}t/2 - 1)^2 + (\sqrt{2}t/2 - 1)^2 - 2}{t}\\ &= \lim_{t\to 0}(t - 2\sqrt{2}) = -2\sqrt{2}\end{aligned}$$

Por otra parte, el plano $y = x$ es el plano perpendicular al plano xy, que contiene al vector v, (el cual corresponde al valor de $\theta = \pi/4$ en la ecuación $x \operatorname{sen}\theta - y\cos\theta = 0$). La intersección de este plano con la superficie $z = (x-1)^2 + (y-1)^2$ es $z = 2(x-1)^2 = \varphi(x)$. Esta función $z = \varphi(x)$ mide *la variación de* z *respecto de* x en la curva de intersección de la superficie con el plano $y = x$. Para obtener la variación de z con respecto a la variable s que se encuentra en el eje marcado por el vector

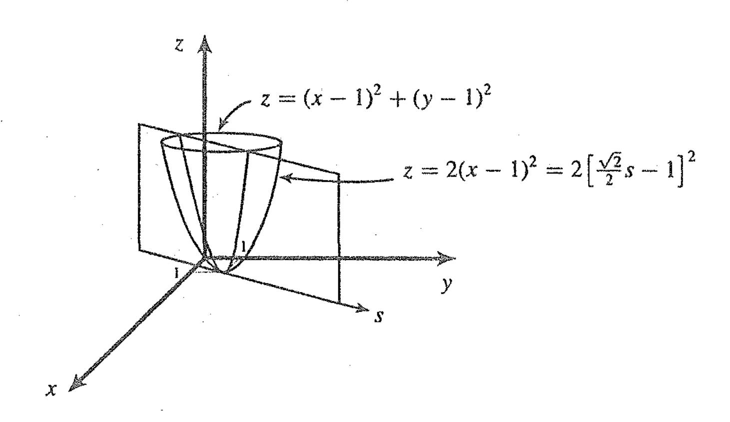


Figura 5. Gráfica del ejemplo 4.

v —que llamamos "eje s"—(es decir, para obtener la "ecuación natural" de la curva en el plano en el que se encuentra ella), observamos que una unidad en el eje x corresponde a $\sqrt{2}$ unidades en el eje s, de modo que la expresión que relaciona este "cambio de unidades" en los ejes x y s es $s\sqrt{2}x$ o bien $x = \sqrt{2}s/2$. Entonces, al sustituir en $z = \varphi(x)$, obtenemos $z = \psi(s) = 2(\sqrt{2}s/2 - 1)^2$ que es la ecuación de la curva en el plano en el que ella se encuentra. Su derivada en $s = 0$ nos da la pendiente de la recta tangente a la curva en ese punto. Tenemos $\psi'(s) = 4(\sqrt{2}s/2 - 1)(\sqrt{2}/2) = 2\sqrt{2}(\sqrt{2}s/2 - 1)$, de donde $\psi'(0) = -2\sqrt{2}$, valor que coincide con el obtenido de la derivada direccional calculado al principio del ejemplo (Fig 5). ■

Retomemos, por último, nuestra discusión sobre el concepto de diferenciabilidad para una función de varias variables. Por supuesto que con la ayuda de las derivadas direccionales podemos formular la siguiente definición más fuerte que la establecida en la sección 2.4: "una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en el punto $\mathbf{x}_0 \in U$ si las derivadas direccionales $\frac{\partial f}{\partial \mathbf{v}}('x_0)$ existen para todo vector $\mathbf{v} \in \mathbb{R}^n$". Observamos que a la luz de esta definición podemos "justificar" por qué la función del ejemplo 6 de la sección 2.4, (la cual no nos conviene que sea diferenciable, pues en tal caso tendría que ser continua en $(0, 0)$) no es diferenciable en $(0, 0)$. En el ejemplo 2 vimos que sus derivadas direccionales no existen en el origen (excepto en las direcciones canónicas). Así pues, a la luz de esta nueva definición, tal función no es diferenciable en $(0, 0)$. Sin embargo, el siguiente ejemplo nos hará desistir de aceptar esta nueva definición como la de diferenciabilidad que estamos procurando.

Ejemplo 5. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

Sea $\mathbf{v} = (a, b) \in \mathbb{R}^2$ un vector unitario dado. Calculemos $\dfrac{\partial f}{\partial \mathbf{v}}(0, 0)$.

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{v}}(0, 0) &= \lim_{t \to 0} \frac{f(ta, tb) - f(0, 0)}{t} \\ &= \lim_{t \to 0} \frac{\dfrac{(ta)^2(tb)}{(ta)^4 + (tb)^2}}{t} \\ &= \lim_{t \to 0} \frac{t^3 a^2 b}{t^3(t^2 a^4 + b^2)} = \frac{a^2}{b} \end{aligned}$$

(Se acepta que $b \neq 0$, pues caso contrario la derivada sería cero —¿por qué?). Entonces esta función posee derivadas direccionales en $(0, 0)$ en todas direcciones. Según lo establecido en la definición anterior esta función debería ser diferenciable en $(0, 0)$. Sin embargo, podemos ver que esta función no es (ni siquiera) continua en $(0, 0)$. En efecto

$$\lim_{(x,y) \to (0,0)} \frac{x^2 y}{x^4 + y^2} = \begin{cases} 0 & (\text{con } y = x) \\ 1/2 & (\text{con } y = x^2) \end{cases}$$

de modo que el límite no existe y la función es por lo tanto discontinua en $(0, 0)$. ■

En la siguiente sección estudiaremos el concepto de diferenciabilidad de una función de varias variables (¡por fin!).

Apéndice El teorema del valor medio

Para funciones de una variable $f: [a, b] \to \mathbb{R}$ tenemos un resultado según el cual si la función es continua en el intervalo cerrado $[a, b]$ y diferenciable en el intervalo abierto (a, b), entonces hay un punto $c \in (a, b)$ en el que

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

lo cual significa geométricamente que (bajo las hipótesis mencionadas para f) siempre tenemos un punto $x = c$ entre $x = a$ y $x = b$, el cual podemos escribir como $c = a + \theta(b - a)$ con $0 < \theta < 1$, en el que la recta tangente a la gráfica de la función $y = f(x)$ (de pendiente $f'(c) = f'(a + \theta(b - a))$) es paralela a la recta secante de la gráfica de f que pasa por los puntos $(a, f(a))$ y $(b, f(b))$ (la cual tiene pendiente $\frac{f(b)-f(a)}{b-a}$). Este resultado se conoce como el teorema del valor medio (o teorema de Lagrange). El teorema se puede reenunciar de la siguiente manera: si la función f es continua en el cerrado $[x_0, x_0 + h]$ y derivable en el abierto $(x_0, x_0 + h)$, entonces hay un $\theta \in (0, 1)$ que

$$f(x_0 + h) - f(x_0) = f'(x_0 + \theta h)h$$

Veremos ahora un resultado análogo para funciones de varias variables en el cual intervienen las derivadas direccionales de la función estudiadas en esta sección. Si $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^n$, introducimos la notación $[\mathbf{x}_0, \mathbf{y}_0]$ para denotar al segmento de recta que une $\mathbf{x}_0$ con $\mathbf{y}_0$. Es decir

$$[\mathbf{x}_0, \mathbf{y}_0] = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \lambda\mathbf{y}_0 - (1 - \lambda)\mathbf{x}_0, \ 0 \leq \lambda \leq 1\}$$

Análogamente, escribiremos

$$(\mathbf{x}_0, \mathbf{y}_0) = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \lambda\mathbf{y}_0 - (1 - \lambda)\mathbf{x}_0, \ 0 < \lambda < 1\}$$

Nótese que $[\mathbf{x}_0, \mathbf{y}_0] = (\mathbf{x}_0, \mathbf{y}_0) \cup \{\mathbf{x}_0, \mathbf{y}_0\}$.

Teorema (del valor medio para funciones de varias variables). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Si $\mathbf{x}_0, \mathbf{y}_0 \in U$, se pide que el conjunto U sea tal que $[\mathbf{x}_0, \mathbf{y}_0] \subset U$. Sea $\mathbf{u}$ un vector unitario en la dirección del vector $\mathbf{y}_0 - \mathbf{x}_0$. Si la función f es continua en los puntos del segmento $[\mathbf{x}_0, \mathbf{y}_0]$ y tiene derivadas direccionales en la dirección del vector $\mathbf{u}$ en los puntos del segmento $(\mathbf{x}_0, \mathbf{y}_0)$, entonces existe un número θ, $0 < \theta < 1$, tal que

$$f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0) = \frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0 + \theta h\mathbf{u})h$$

(donde $h = \|\mathbf{y}_0 - \mathbf{x}_0\|$).

Demostración. Considere la función $\phi: [0, h] \to \mathbb{R}$ dada por

$$\phi(t) = f(\mathbf{x}_0 + t\mathbf{u})$$

Ciertamente la función ϕ es continua en $[0, h]$, pues f lo es en $[\mathbf{x}_0, \mathbf{y}_0]$. Además obsérvese que

$$\begin{aligned}\phi'(t) &= \lim_{\bar{h} \to 0} \frac{\phi(t + \bar{h}) - \phi(t)}{\bar{h}} = \lim_{\bar{h} \to 0} \frac{f(\mathbf{x}_0 + (t + \bar{h})\mathbf{u}) - f(\mathbf{x}_0 + t\mathbf{u})}{\bar{h}} \\ &= \lim_{\bar{h} \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{u} + \bar{h}\mathbf{u}) - f(\mathbf{x}_0 + t\mathbf{u})}{\bar{h}} = \frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0 + t\mathbf{u})\end{aligned}$$

de modo que para $t \in (0, h)$, $\phi'(t)$ existe y es la derivada direccional de la función f en el punto $\mathbf{x}_0 + t\mathbf{u} \in (\mathbf{x}_0, \mathbf{y}_0)$ en la dirección del vector $\mathbf{u}$. Aplicando entonces el teorema del valor medio a la función ϕ, concluimos que existe un número $\theta \in (0, 1)$ que da

$$\phi(h) - \phi(0) = \phi'(\theta h)h$$

es decir, de modo que

$$f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0) = \frac{\partial f}{\partial u}(\mathbf{x}_0 + \theta h\mathbf{u})h$$

tal como se quería demostrar. Q.E.D.

Ejercicios (Capítulo 2, Sección 5)

En los ejercicios 1–6, identifique las expresiones dadas como derivadas direccionales de funciones de varias variables en la dirección de un vector unitario v. Obtenga la derivada direccional que se indica.

1. $\displaystyle \lim_{t\to 0} \frac{(x + t/\sqrt{2})(y + t/\sqrt{2})^2 - xy^2}{t}$

2. $\displaystyle \lim_{t\to 0} \frac{(x - \sqrt{3}t/2)^{1/2}(y + t/2)^{1/2} - \sqrt{xy}}{t}$

3. $\displaystyle \lim_{t\to 0} \frac{\ln \operatorname{sen}^2((x + t)^4 y) - \ln \operatorname{sen}^2(x^4 y)}{t}$

4. $\displaystyle \lim_{t\to 0} \frac{(y + t)^2 \cos^3(xy + xt) - y^2 \cos^3(xy)}{t}$

5. $\displaystyle \lim_{t\to 0} \frac{(x + 2t/3)(y + 2t/3)(z - t/3) - xyz}{t}$

6. $\displaystyle \lim_{t\to 0} \frac{x^2(y - \sqrt{3}t/2)(z - t/2) - x^2yz}{t}$

En los ejercicio 7–15, calcule la derivada direccional de la función dada en la dirección del vector indicado.

7. $f(x, y) = 3x - 2y$, $\mathbf{v} = (1/\sqrt{2}, 1/\sqrt{2})$
8. $f(x, y) = 3x - 2y$, $\mathbf{v} = (-1/\sqrt{2}, -1/\sqrt{2})$
9. $f(x, y) = x^2 + y^2$, $\mathbf{v} = (a, b)$, en el punto (0,0).
10. $f(x, y) = xy^2 + x^2y$, $\mathbf{v} = (1, 0)$.
11. $f(x, y) = x^3\sqrt{1 + 3\tan^6(x^2 + x^{102})}$, $\mathbf{v} = (0, 1)$.
12. $f(x, y, z) = z \operatorname{sen} y^3 \cos(x^5 + \tan y^3)$, $\mathbf{v} = (0, 0, 1)$.
13. $f(x, y, z) = xyz$, $\mathbf{v} = \left(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right)$.
14. $f(x, y, z, u) = xyzu$, $\mathbf{v} = \left(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}, 0\right)$.

15. $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$, $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$.

16. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $\mathbf{u} \in \mathbb{R}^n$ un vector no nulo de $\mathbb{R}^n$, no necesariamente de norma 1 y sea $\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$. Defina la derivada de f en el punto $\mathbf{p} \in U$ en la dirección del vector $\mathbf{u} \in \mathbb{R}^n$ de manera análoga a como se hizo en el libro (para el caso de vectores unitarios)

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{p}) = \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{t}$$

cuando este límite existe. Demuestre que

$$\frac{\partial f}{\partial \mathbf{v}} = \frac{1}{\|\mathbf{u}\|} \frac{\partial f}{\partial \mathbf{u}}$$

Verifique este resultado con la función $f(x, y) = x^2 + y^2$, y el vector $\mathbf{u} = (1, 1)$.

17. Con el resultado del ejercicio anterior, calcule la derivada direccional de la función $f(x, y) = 3x + 2y + 7z$ en la dirección del vector $\mathbf{u} = (3, 2, -5)$.

18. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$. Sea $\mathbf{v} \in \mathbb{R}^2$ un vector unitario dado. Demuestre que si la derivada direccional de f en la dirección de $\mathbf{v}$ existe en el punto $\mathbf{p} \in U$, entonces f es continua en $\mathbf{p}$ en la dirección del vector $\mathbf{v}$ (ver ejercicio 71 de la sección 3). ¿Es verdadera la afirmación recíproca?

19. Sean $f, g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ dos funciones definidas en el conjunto abierto U de $\mathbb{R}^2$ con derivadas direccionales en $\mathbf{p} \in U$ en la dirección de un vector unitario $\mathbf{v} \in \mathbb{R}^2$ dado.

   a. Demuestre que la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F = f + g$ tiene derivada direccional en $\mathbf{p}$ en la dirección de $\mathbf{v}$ y que ésta es

$$\frac{\partial F}{\partial \mathbf{v}} = \frac{\partial f}{\partial \mathbf{v}} + \frac{\partial g}{\partial \mathbf{v}}$$

   b. Demuestre que la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F = fg$ tiene derivada direccional en $\mathbf{p}$ en la dirección de $\mathbf{v}$ y que ésta es

$$\frac{\partial F}{\partial \mathbf{v}} = f\frac{\partial g}{\partial \mathbf{v}} + g\frac{\partial f}{\partial \mathbf{v}}$$

   c. Demuestre que si $g(x, y) \neq 0$ en U, la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $F = \frac{f}{g}$ tiene derivada direccional en $\mathbf{p}$ en la dirección de $\mathbf{v}$ y que ésta es

$$\frac{\partial F}{\partial \mathbf{v}} = \frac{g\frac{\partial f}{\partial \mathbf{v}} - f\frac{\partial g}{\partial \mathbf{v}}}{g^2}$$

20. Compruebe que las derivadas direccionales obtenidas en los ejercicios 7–11 se pueden escribir como

$$\frac{\partial f}{\partial \mathbf{v}} = a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial y}$$

en que $\mathbf{v} = (a, b)$ es el vector dado en la dirección del cual se calculó la derivada.

21. Compruebe que las derivadas direccionales obtenidas en los ejercicios 12–13 se pueden escribir como

$$\frac{\partial f}{\partial \mathbf{v}} = a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial y} + c\frac{\partial f}{\partial z}$$

donde $\mathbf{v} = (a, b, c)$ es el vector dado en la dirección de la cual se calculó la derivada.

22. Verifique que la derivada direccional de la función del ejercicio 15 se puede escribir como

$$\frac{\partial f}{\partial \mathbf{v}} = \alpha_1\frac{\partial f}{\partial x_1} + \alpha_2\frac{\partial f}{\partial x_2} + \cdots + \alpha_n\frac{\partial f}{\partial x_n}$$

23. Los resultados de los 3 ejercicios anteriores sugieren que la fórmula

$$\frac{\partial f}{\partial \mathbf{v}} = \alpha_1\frac{\partial f}{\partial x_1} + \alpha_2\frac{\partial f}{\partial x_2} + \cdots + \alpha_n\frac{\partial f}{\partial x_n}$$

se puede usar para calcular la derivada direccional de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ en la dirección del vector $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. En efecto, esto es efectivamente cierto para "algunas" funciones que empezaremos a estudiar en la siguiente sección. Para convencernos que tal fórmula no siempre es cierta, considere la función $f: \mathbb{R}^2 \to \mathbb{R}$ del ejemplo 5

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

Demuestre que

$$\frac{\partial f}{\partial \mathbf{v}}(0, 0) \neq a\frac{\partial f}{\partial x}(0, 0) + b\frac{\partial f}{\partial y}(0, 0)$$

donde $\mathbf{v} = (a, b)$ es un vector dado.

24. **a.** Sea $\mathbf{v} = (a, b)$ un vector unitario en $\mathbb{R}^2$. Demuestre que el plano perpendicular a $z = 0$, que pasa por (x_0, y_0), para el que $\mathbf{v}$ es un vector paralelo (decimos que el plano está en la dirección del vector $\mathbf{v}$), se puede escribir como

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \end{cases} \qquad t \in \mathbb{R}$$

b. Demuestre que la curva de intersección de la superficie $z = f(x, y)$ con un plano como el descrito en el inciso anterior es $z = f(x_0 + at, y_0 + bt)$.

c. Considere la función $\phi: I \subseteq \mathbb{R} \to \mathbb{R}$, $\phi(t) = f(x_0 + at, y_0 + bt)$, cuyo dominio es el conjunto de la recta dado por $I = \{t \in \mathbb{R} | (x_0 + at, y_0 + bt) \in \text{ dominio de } f\}$. Nótese que el conjunto de imágenes de ϕ es justamente la curva del inciso anterior. Demuestre que $\phi'(0)$ existe si y sólo si se da la derivada de f en (x_0, y_0) en la dirección de $\mathbf{v} = (a, b)$ existe. En tal caso $\phi'(0) = \frac{\partial f}{\partial \mathbf{v}}(x_0, y_0)$.

d. Demuestre que la ecuación de la recta tangente (cuando existe) a la curva del inciso **b**, en el punto $(x_0, y_0, f(x_0, y_0))$ es

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = f(x_0, y_0) + \dfrac{\partial f}{\partial \mathbf{v}}(x_0, y_0)t \end{cases} \qquad t \in \mathbb{R}$$

e. Obtenga, como caso particular de los resultados de los incisos anteriores, los resultados de los problemas 39 y 40 de la sección anterior.

25. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función continua tal que sus derivadas parciales existan en todo punto $\mathbf{p} \in \mathbb{R}^2$. Demuestre que existen $\theta_1, \theta_2 \in (0, 1)$ tales que

$$f(x_0 + h, y_0) - f(x_0, y_0) = h\frac{\partial f}{\partial x}(x_0 + \theta_1 h, y_0)$$

$$f(x_0, y_0 + h) - f(x_0, y_0) = h\frac{\partial f}{\partial y}(x_0, y_0 + \theta_2 h)$$

2.6 Diferenciabilidad

Recordemos de nuevo que una función de una variable $f: I \subseteq \mathbb{R} \to \mathbb{R}$ se dice que es diferenciable en $x_0 \in I$ si existe el límite

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

En tal caso al valor del límite se le llama "derivada de f en x_0" y se denota por $f'(x_0)$. Por supuesto que un primer intento para conseguir un concepto equivalente para funciones de varias variables, digamos $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ sería *copiar* la definición anterior extendiéndola a esta nueva situación. Esto, sin embargo, conduce a una expresión que carece de sentido. En efecto, si f está definida en un subconjunto U de $\mathbb{R}^n$, $n > 1$, tanto x_0 como h en la definición anterior serían vectores de $\mathbb{R}^n$, y así, en la expresión que define a $f'(x_0)$ aparece *una división por* h, operación que carece de sentido con vectores de $\mathbb{R}^n$, $n > 1$. A pesar de este contratiempo, nuestro interés seguirá siendo tratar de "copiar la esencia" de la definición de diferenciabilidad que conocemos para plasmarla en el concepto que queremos establecer de diferenciabilidad para funciones de varias variables. Lo que haremos ahora será replantear la definición para una variable, de modo que en ella no aparezcan divisiones (que es la operación que no se puede copiar para vectores de $\mathbb{R}^n$ con $n \geq 2$).

Una manera equivalente —y más elegante— de establecer el concepto de diferenciabilidad para funciones de una variable es la siguiente: la función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ es diferenciable en $x_0 \in I$ si existe una constante A tal que

$$f(x_0 + h) = f(x_0) + Ah + r(h)$$

donde

$$\lim_{h \to 0} \frac{r(h)}{h} = 0$$

En efecto, escribiendo explícitamente la constante A de la última definición, tenemos

$$A = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{r(h)}{h}$$

de donde (al tomar límite cuando $h \to 0$) se puede ver fácilmente la equivalencia entre las definiciones (poniendo $A = f'(x_0)$).

Más aún, es interesante la interpretación geométrica de la nueva definición, como se ve en la figura 1

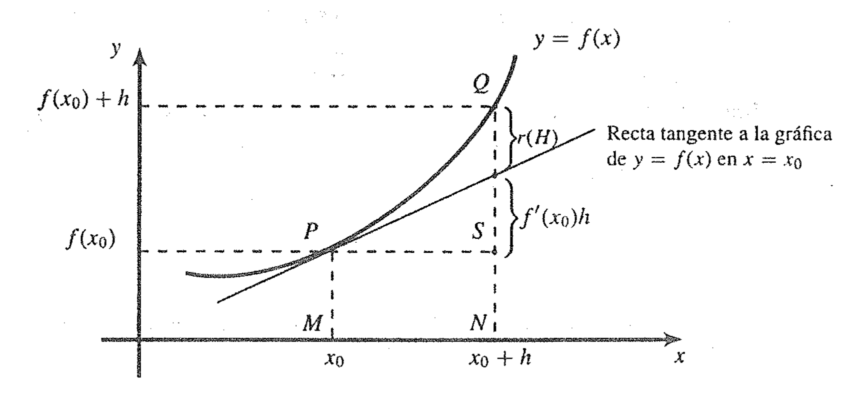


Figura 1. La derivada de una función.

La ecuación de la recta tangente a la gráfica de la función $y = f(x)$ en el punto $P = (x_0, f(x_0))$ es, como sabemos

$$y = f'(x_0)(x - x_0) + f(x_0)$$

de manera que el punto R de abscisa $x_R = x_0 + h$ tiene por ordenada $y_R = f'(x_0)h + f(x_0)$, de donde se ve fácilmente que la magnitud del segmento $\overline{RS}$ es $f'(x_0)h$ y por tanto, con $r(h) = f(x_0 + h) - f(x_0) - f'(x_0)h$, vemos que $r(h)$ es justamente la distancia entre la curva y la recta tangente en el punto P. De la figura también se desprende que a medida que h tiende a cero, el residuo $r(h)$ (llamado así, pensando en que es "lo que le falta a la recta tangente para describir a la curva $y = f(x)$ en los alrededores de P") también tiende a cero. Sin embargo, esto no es el punto importante en la definición de diferenciabilidad, pues este hecho (el que $\lim_{h\to 0} r(h) = 0$) *lo único que nos dice es que la función es continua en* P (¿por qué?). Lo importante, cuando se estudia la diferenciabilidad de funciones, es que el residuo $r(h)$ tiende a cero *más rápido* de lo que h lo hace. Esta condición está expresada en la parte de la definición que dice que

$$\lim_{h\to 0} \frac{r(h)}{h} = 0$$

Intuitivamente, podemos pensar que la única manera como se puede dar esta condición del residuo, es cuando la recta tangente se *"embarra"* con la curva en los alrededores de P. En otras palabras, la curva tiene que ser "suave" en P, para que "se pueda ver localmente como una recta" (su recta tangente). Esta es en realidad la idea geométrica tras el concepto de diferenciabilidad de una función en un punto P: el buen comportamiento de la función en torno al punto P para que en se pueda aproximar por un comportamiento lineal (el de su recta tangente).

Consideremos ahora una función de dos variables $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$. Sea $\mathbf{p} = (x_0, y_0) \in U$. Generalizando la definición de diferenciabilidad para funciones de una variable estudiada anteriormente, diremos que la función f es *diferenciable* en el punto $\mathbf{p} = (x_0, y_0)$ si hay constantes A_1 y A_2 de modo que

$$f((x_0, y_0) + (h_1, h_2)) = f(x_0, y_0) + A_1h_1 + A_2h_2 + r(h_1, h_2) \qquad \text{D}$$

donde

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1,h_2)}{\|(h_1,h_2)\|} \tag{R}$$

No se puede negar que esta definición se ve como una generalización natural de la dada para funciones de una variable: el incremento h ha sido sustituido por el vector $\mathbf{h} = (h_1, h_2) \in \mathbb{R}^2$, y al residuo r que depende de h, se le ha pedido que, al ser dividido por la norma (en este caso no podemos dividir simplemente por h) de $\mathbf{h}$, tienda a cero cuando h tiende a cero.

Supongamos que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en $\mathbf{p} = (x_0, y_0) \in U$. Veamos cuáles tienen que ser las constantes A_1 y A_2 de la definición anterior. Poniendo $\mathbf{h} = (h_1, 0)$ en la expresión (D), podemos escribir

$$\frac{r(h_1,0)}{h_1} = \frac{r(h_1)}{h_1} = \frac{f(x_0+h_1,y_0)-f(x_0,y_0)}{h_1} - A_1$$

Si tomamos límite cuando $h_1 \to 0$ (i.e. cuando $\mathbf{h} \to \mathbf{0}$) y usamos la propiedad (R) del residuo en la definición, obtenemos

$$\lim_{h_1\to 0} \frac{r(h_1)}{h_1} = \lim_{h_1\to 0}\left(\frac{f(x_0+h_1,y_0)-f(x_0,y_0)}{h_1} - A_1\right) = 0$$

de donde vemos que

$$A_1 = \lim_{h_1\to 0} \frac{f(x_0+h_1,y_0)-f(x_0,y_0)}{h_1} = \frac{\partial f}{\partial x}(x_0,y_0)$$

Un argumento análogo (poniendo $\mathbf{h} = (0, h_2)$) nos conduce a

$$A_2 = \lim_{h_2\to 0} \frac{f(x_0,y_0+h_2)-f(x_0,y_0)}{h_2} = \frac{\partial f}{\partial y}(x_0,y_0)$$

Se observa entonces que una condición *necesaria* para que una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ sea diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in U$ es que existan sus derivadas parciales en ese punto. Sin embargo, tal condición está muy lejos de ser suficiente como se verá posteriormente.

Podemos entonces, en resumen, establecer la definición siguiente

Definición. Se dice que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$, es *diferenciable* en el punto $\mathbf{p} = (x_0, y_0) \in U$, si se dan las derivadas parciales de f en $\mathbf{p}$

$$A_1 = \frac{\partial f}{\partial x}(x_0,y_0), \qquad A_2 = \frac{\partial f}{\partial y}(x_0,y_0)$$

y si el residuo $r(h_1, h_2)$ definido en la expresión

$$f((x_0,y_0)+(h_1,h_2)) = f(x_0,y_0) + A_1h_1 + A_2h_2 + r(h_1,h_2)$$

tiene la propiedad

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1,h_2)}{\|(h_1,h_2)\|} = 0$$ ■

Veamos ahora que esta definición sí garantiza la continuidad de la función en el punto $\mathbf{p}$ en donde es diferenciable (como acontece con funciones de una variable).

Teorema 2.6.1 Si la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$, es diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in U$, entonces es continua en ese punto.

Demostración. Siendo f diferenciable en $\mathbf{p} = (x_0, y_0)$ se tiene

$$f((x_0, y_0) + (h_1, h_2)) = f(x_0, y_0) + A_1 h_1 + A_2 h_2 + r(h_1, h_2)$$

Tomando el límite cuando $(h_1, h_2) \to \mathbf{0}$ y observando que de la condición establecida en la definición para el residuo $r(h_1, h_2)$, se deduce que

$$\lim_{(h_1,h_2)\to(0,0)} r(h_1, h_2) = 0$$

vemos que

$$\lim_{(h_1,h_2)\to(0,0)} f\big((x_0, y_0) + (h_1, h_2)\big) = \lim_{(h_1,h_2)\to(0,0)} \big(f(x_0, y_0) + A_1 h_1 + A_2 h_2 + r(h_1, h_2)\big) = f(x_0, y_0)$$

lo cual significa precisamente que la función f es continua en (x_0, y_0) Q.E.D.

Veamos ahora algunos ejemplos.

Ejemplo 1. La función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

(ver ejemplo 6 sección 4) no puede ser diferenciable en el origen, pues aunque las dos derivadas parciales de la función existen en ese punto (son ambas iguales a cero), ésta es discontinua en ese punto. Por el teorema anterior, la función no es diferenciable en (0, 0), de lo cual también podemos convencernos directamente de la definición, escribiendo

$$f\big((0, 0) + (h_1, h_2)\big) = f(0, 0) + (0)(h_1) + (0)h_2 + r(h_1, h_2)$$

de donde

$$r(h_1, h_2) = \frac{h_1 h_2}{h_1^2 + h_2^2}$$

y

$$\frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = \frac{r(h_1, h_2)}{\sqrt{h_1^2 + h_2^2}} = \frac{h_1 h_2}{\big(h_1^2 + h_2^2\big)^{3/2}}$$

Tomando límite cuando $(h_1, h_2) \to (0, 0)$ obtenemos

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = \lim_{(h_1,h_2)\to(0,0)} \frac{h_1 h_2}{\big(h_1^2 + h_2^2\big)^{3/2}}$$

$$\overset{h_1 = h_2}{=\!=\!\longrightarrow}$$

$$= \lim_{h_1\to 0} \frac{h_1^2}{2^{3/2} h_1^3} = \lim_{h_1\to 0} \frac{1}{2^{3/2} h_1}$$

$|h_1|^3$

límite que no existe. ▪

Ejemplo 2. Una función polinomial $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) \sum_{i,j=0}^{n} a_{ij} x^i y^j$ es diferenciable en todo punto $(x_0, y_0) \in \mathbb{R}^2$. Este es un hecho de verificación directa de la definición, que requiere de algunas cuentas que dejamos al lector. Por simplicidad, verifiquemos esta afirmación con la función $f(x, y) = xy^2$ en el origen. Escribiendo

$$f\big((0, 0) + (h_1, h_2)\big) = f(0, 0) + \frac{\partial f}{\partial x}(0, 0)h_1 + \frac{\partial f}{\partial y}(0, 0)h_2 + r(h_1, h_2)$$

nos queda

$$h_1 h_2^2 = r(h_1, h_2)$$

de donde

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = \lim_{(h_1,h_2)\to(0,0)} \frac{h_1 h_2^2}{\left(h_1^2 + h_2^2\right)^{3/2}}$$

$$\xrightarrow[h_2 = r \operatorname{sen}\theta]{h_1 = r\cos\theta} = \lim_{r\to 0} r^2 \cos\theta \operatorname{sen}^2\theta = 0$$

tal como lo requería la definición.

Verifiquemos también la diferenciabilidad de la función polinomial $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2$ en un punto arbitrario $(x_0, y_0) \in \mathbb{R}^2$. Tenemos

$$f\big((x_0, y_0) + (h_1, h_2)\big) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h_1 + \frac{\partial f}{\partial y}(x_0, y_0)h_2 + r(h_1, h_2)$$

o sea

$$(x_0 + h_1)^2 + (y_0 + h_2)^2 = x_0^2 + y_0^2 + 2x_0 h_1 + 2y_0 h_2 + r(h_1, h_2)$$

de donde $r(h_1, h_2) = h_1^2 + h_2^2$, y entonces

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = \lim_{(h_1,h_2)\to(0,0)} \frac{h_1^2 + h_2^2}{\left(h_1^2 + h_2^2\right)^{1/2}}$$

$$= \lim_{(h_1,h_2)\to(0,0)} \left(h_1^2 + h_2^2\right)^{1/2} = 0$$

▪

Ejemplo 3. La función $f: \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = \sqrt{x^2 + y^2}$ no es diferenciable en el origen. En efecto, las derivadas parciales $\frac{\partial f}{\partial x}(0, 0)$ y $\frac{\partial f}{\partial y}(0, 0)$ serían

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h\to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h\to 0} \frac{|h|}{h}$$

(de igual modo para $\frac{\partial f}{\partial y}(0, 0)$) que es un límite no existente (pues $\lim_{h\to 0^+} \frac{|h|}{h} = 1$ y $\lim_{h\to 0^-} \frac{|h|}{h} = -1$). Entonces, siendo la existencia de las derivadas parciales en el punto $\mathbf{p}$ una condición necesaria para la diferenciabilidad de la función en $\mathbf{p}$, concluimos que la función $f(x, y) = \sqrt{x^2 + y^2}$ no es diferenciable en el origen. ▪

Después de ver estos ejemplos, se podría presentar un panorama desconsolador si se piensa en la diferenciabilidad de una función de varias variables como una propiedad muy restrictiva de difícil

verificación. Esto podría, en algún sentido, ser cierto. Sin embargo, nuestro objetivo será poder desarrollar cierta "intuición en torno a la diferenciabilidad de funciones de varias variables" que nos permita "detectar a priori" las funciones *no* diferenciables. Esta intuición tiene que ver en parte con los siguientes teoremas que estudiarenos y en parte con la idea geométrica de la diferenciabilidad que veremos en la sección 10. Debemos, sin embargo, decir que las funciones con las que trabajaremos en este libro serán, a menos que se enfatize lo contrario, diferenciables en todo su dominio.

Como era de esperarse, la suma, el producto y el cociente de funciones diferenciables son también una función diferenciable (como ocurre con las funciones de una sola variable). He aquí el contenido del siguiente teorema.

Teorema 2.6.2 Sean $f, g\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones definidas en el conjunto abierto U de $\mathbb{R}^n$, diferenciables en $\mathbf{p} \in U$. Entonces:

a. la suma $f + g\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, $(f+g)(\mathbf{p}) = f(\mathbf{p}) + g(\mathbf{p})$ es una función diferenciable en $\mathbf{p}$.

b. el producto $fg\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, $(fg)(\mathbf{p}) = f(\mathbf{p})g(\mathbf{p})$ es una función diferenciable en $\mathbf{p}$.

c. si $g(\mathbf{p}) \neq 0$, el cociente $\frac{f}{g}\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, $\left(\frac{f}{g}\right)(\mathbf{p}) = \frac{f(\mathbf{p})}{g(\mathbf{p})}$ es diferenciable en $\mathbf{p}$.

Demostración. (opcional) La demostración de los 3 incisos es similar. Presentaremos sólo la del inciso **b** en el caso $n = 2$. Escribamos el punto $\mathbf{p}$ como $\mathbf{p} = (x_0, y_0)$. Debemos probar que el residuo $r(h_1, h_2)$ definido por la expresión

$$(fg)\big((x_0, y_0) + (h_1, h_2)\big) = (fg)(x_0, y_0) + \frac{\partial}{\partial x}(fg)(x_0, y_0)h_1 + \frac{\partial}{\partial y}(fg)(x_0, y_0)h_2 + r(h_1, h_2)$$

tiene la propiedad

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = 0$$

Si vemos que (ver ejercicio 37 de la sección 4)

$$\frac{\partial}{\partial x}(fg) = f\frac{\partial g}{\partial x} + g\frac{\partial f}{\partial x}, \qquad \frac{\partial}{\partial y}(fg) = f\frac{\partial g}{\partial y} + g\frac{\partial f}{\partial y}$$

nos queda, escribiendo $\mathbf{p} = (x_0, y_0)$, $\mathbf{h} = (h_1, h_2)$

$$(fg)(\mathbf{p}+\mathbf{h}) = (fg)(\mathbf{p}) + \left(f(\mathbf{p})\frac{\partial g}{\partial x}(\mathbf{p}) + g(\mathbf{p})\frac{\partial f}{\partial x}(\mathbf{p})\right)h_1 + \left(f(\mathbf{p})\frac{\partial g}{\partial y}(\mathbf{p}) + g(\mathbf{p})\frac{\partial f}{\partial y}(\mathbf{p})\right)h_2 + r(h_1, h_2)$$

expresión que podemos reescribir como

$$\begin{aligned} r(h_1, h_2) &= g(\mathbf{p}+\mathbf{h})\big(f(\mathbf{p}+\mathbf{h}) - f(\mathbf{p}) - \frac{\partial f}{\partial x}(\mathbf{p})h_1 - \frac{\partial f}{\partial y}(\mathbf{p})h_2\big) \\ &\quad + f(\mathbf{p})\left(g(\mathbf{p}+\mathbf{h}) - g(\mathbf{p}) - \frac{\partial g}{\partial x}(\mathbf{p})h_1 - \frac{\partial g}{\partial y}(\mathbf{p})h_2\right) \\ &\quad + \frac{\partial f}{\partial x}(\mathbf{p})\big(g(\mathbf{p}+\mathbf{h}) - g(\mathbf{p})\big)h_1 + \frac{\partial f}{\partial y}(\mathbf{p})\big(g(\mathbf{p}+\mathbf{h}) - g(\mathbf{p})\big)h_2 \end{aligned}$$

de donde

$$\frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = g(\mathbf{p} + \mathbf{h}) \frac{f(\mathbf{p} + \mathbf{h}) - f(\mathbf{p}) - \frac{\partial f}{\partial x}(\mathbf{p})h_1 - \frac{\partial f}{\partial y}(\mathbf{p})h_2}{\|(h_1, h_2)\|}$$
$$+ f(\mathbf{p}) \frac{g(\mathbf{p} + \mathbf{h}) - g(\mathbf{p}) - \frac{\partial g}{\partial x}(\mathbf{p})h_1 - \frac{\partial g}{\partial y}(\mathbf{p})h_2}{\|(h_1, h_2)\|}$$
$$+ \frac{\partial f}{\partial x}(\mathbf{p})\big(g(\mathbf{p} + \mathbf{h}) - g(\mathbf{p})\frac{h_1}{\|(h_1, h_2)\|} + \frac{\partial f}{\partial y}(\mathbf{p})\big(g(\mathbf{p} + h) - g(\mathbf{p})\big)\frac{h_2}{\|(h_1, h_2)\|}$$

Al tomar el límite cuando $(h_1, h_2) \to (0, 0)$ vemos que los dos primeros sumandos tienden a cero pues las funciones f y g son, por hipótesis, diferenciables en $\mathbf{p}$ (los segundos factores de cada uno de estos sumandos definen los residuos —divididos por $\|\mathbf{h}\|$— de las funciones f y g respectivamente). Por otra parte, como

$$\frac{|h_1|}{\|(h_1, h_2)\|} \leq 1, \qquad \frac{|h_2|}{\|(h_1, h_2)\|} \leq 1$$

para todo $(h_1, h_2) \in \mathbb{R}^2$ no nulo (es decir, estas expresiones se mantienen acotadas cerca del origen de $\mathbb{R}^2$), y como g es diferenciable en $\mathbf{p}$, es continua en $\mathbf{p}$ (teorema 2.6.1), entonces

$$\lim_{h \to 0}(g(p + h) - g(p)) = 0$$

por lo cual también los dos últimos sumandos tienden a cero cuando $\mathbf{h} \to 0$. Esto muestra lo que queríamos. Q.E.D.

Para el caso de la composición de funciones, también es cierto que si componemos la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ diferenciable en el punto $\mathbf{p} = (x_0, y_0)$, con la función $g: I \subseteq \mathbb{R} \to \mathbb{R}$ (donde I es un conjunto abierto de $\mathbb{R}$ que contiene al rango de f) diferenciable en $f(x_0, y_0) \in I$, entonces la composición $g \circ f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, será diferenciable en $\mathbf{p}$. Este es un punto que se discutirá con más amplitud en el siguiente capítulo.

Ejemplo 4. La función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = \frac{x^2 - y^2 - 1}{x^2 + y^2 + 1}$ es diferenciable en todo el espacio $\mathbb{R}^2$, pues es el cociente de dos funciones polinomiales que, como se vio en el ejemplo 2, siempre son diferenciables (y además la función del denominador siempre es no nula). ■

Ejemplo 5. La función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = e^{-(x^2+y^2)}$ es diferenciable en todo el espacio $\mathbb{R}^2$, pues es la composición de la función $\varphi(x, y) = x^2 + y^2$ que siempre es diferenciable, con la función $\psi(x) = e^{-x}$ que también lo es. ■

Ejemplo 6. La función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = x \operatorname{sen} \frac{x + y}{1 + x^2} + (y^2 - xy + 1)\cos(x^2 + y^2)$$

es diferenciable en todo (punto de) el espacio $\mathbb{R}^2$, pues está formada por sumandos, productos, cocientes y composiciones de funciones diferenciables. ■

La situación en el caso de funciones de n variables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definidas en el conjunto abierto U de $\mathbb{R}^n$ es completamente análoga: la función será diferenciable en el punto $\mathbf{x}_0 \in U$ si se presentan las derivadas parciales $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$, $i = 1, 2, \ldots, n$ y si el residuo $r(\mathbf{h})$ definido como

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\mathbf{h}_i + r(\mathbf{h})$$

(donde $\mathbf{h} = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ es tal que $\mathbf{x}_0 + \mathbf{h} \in U$) tiene la propiedad de que

$$\lim_{\mathbf{h}\to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

La diferenciabilidad de la función f en $\mathbf{x}_0$ implica su continuidad en ese punto.

Cuando una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto $U \in \mathbb{R}^n$ es diferenciable en todo punto de su dominio, decimos simplemente que es "diferenciable".

Discutamos, por último, qué papel que juegan las derivadas parciales de una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto $\mathbf{x}_0 \in U$ con la diferenciabilidad de la función en este punto. Como hemos visto, el hecho de que haya tales derivadas no garantiza que la función sea diferenciable. Sin embargo, el siguiente teorema nos dice que si las derivadas parciales de la función (existen y) son continuas en el punto $\mathbf{x}_0$, entonces sí podremos concluir la diferenciabilidad de la misma en ese punto.

Cuando decimos que la derivada parcial $\frac{\partial f}{\partial x_i}$ de la función f respecto de su i-ésima variable, es continua en el punto $\mathbf{x}_0$, nos estamos refiriendo a que la función $\frac{\partial f}{\partial x_i}: \tilde{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, definida en $\tilde{U} \subseteq U$,[1] que a cada punto $\mathbf{x}_0 \in \tilde{U}$ le asocia la derivada parcial $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$, es continua en $\mathbf{x}_0$, es decir que

$$\lim_{h\to 0} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + \mathbf{h}) = \frac{\partial f}{\partial x_i}(\mathbf{x}_0)$$

La continuidad de funciones de varias variables es un concepto que fácilmente se maneja de manera intuitiva: nos resulta fácil —en general— decidir a priori cuándo una función es continua o no lo es. Muchas de las funciones con las que trabajamos en la práctica son funciones cuyas derivadas parciales "se ven" como funciones continuas. Con la ayuda del siguiente teorema, podremos concluir entonces que tales funciones son diferenciables.

Teorema 2.6.3 Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Si las funciones (derivadas parciales) $\frac{\partial f}{\partial x_i}: \tilde{U} \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, n$, $\tilde{U} \subseteq U$, son continuas en el punto $\mathbf{x}_0 \in \tilde{U}$, entonces f es diferenciable en $\mathbf{x}_0$.

Demostración. (opcional, caso $n = 2$). Sea entonces la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ tal que las funciones

$$\frac{\partial f}{\partial x}: \tilde{U} \subseteq \mathbb{R}^2 \to \mathbb{R}, \qquad \frac{\partial f}{\partial y}: \tilde{U} \subseteq \mathbb{R}^2 \to \mathbb{R}$$

[1] Esta funciones podrían no estar definidas en todo el conjunto abierto U, pero deben estarlo al menos en una bola abierta con centro en $\mathbf{x}_0$ —contenida en U—, denotada por $\tilde{U}$. Por ejemplo, considere la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = x^{1/3} + y^{1/3}$. Sus derivadas parciales son $\frac{\partial f}{\partial x} = \frac{1}{3}x^{-1/3}$, $\frac{\partial f}{\partial y} = \frac{1}{3}y^{-1/3}$. Aunque f está definida en todo $\mathbb{R}^2$, sus parciales no lo están.

sean continuas en el punto $(x_0, y_0) \in \tilde{U}$. Consideremos el residuo

$$r(h_1, h_2) = f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h_1 - \frac{\partial f}{\partial y}(x_0, y_0)h_2$$

Queremos ver que

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = 0$$

Escribamos la expresión del residuo como

$$\begin{aligned} r(h_1, h_2) = f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0 + h_2) + \\ f(x_0, y_0 + h_2) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h_1 - \frac{\partial f}{\partial y}(x_0, y_0)h_2 \end{aligned}$$

Si consideramos la función $\varphi(x) = f(x, y_0+h_2)$ (que es diferenciable en x_0 —¿por qué?) y aplicamos el Teorema del Valor Medio (Lagrange) en el intervalo $[x_0, x_0+h_1]$ obtenemos que existe un número $\theta_1 \in [0, 1]$ tal que

$$\varphi(x_0 + h_1) - \varphi(x_0) = \varphi'(x_0 + \theta_1 h_1)h_1$$

o sea

$$f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0 + h_2) = \frac{\partial f}{\partial x}(x_0 + \theta_1 h_1, y_0 + h_2)h_1$$

Del mismo modo, al aplicar el Teorema del Valor Medio en la función $\psi(y) = f(x_0, y)$ en el intervalo $[y_0, y_0 + h_2]$ obtenemos que

$$f(x_0, y_0 + h_2) - f(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0 + \theta_2 h_2)h_2$$

donde θ_2 es un número entre 0 y 1. Entonces

$$\begin{aligned} r(h_1, h_2) &= \frac{\partial f}{\partial x}(x_0 + \theta_1 h_1, y_0 + h_2) + \frac{\partial f}{\partial y}(x_0, y_0 + \theta_2 h_2)h_2 - \frac{\partial f}{\partial x}(x_0, y_0)h_1 - \frac{\partial f}{\partial y}(x_0, y_0)h_2 \\ &= \left[\frac{\partial f}{\partial x}(x_0 + \theta_1 h_1, y_0 + h_2) - \frac{\partial f}{\partial x}(x_0, y_0)\right] h_1 + \left[\frac{\partial f}{\partial y}(x_0, y_0 + \theta_2 h_2) - \frac{\partial f}{\partial y}(x_0, y_0)\right] h_2 \end{aligned}$$

Al dividir por la norma del vector (h_1, h_2) obtenemos

$$\begin{aligned} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} &= \left[\frac{\partial f}{\partial x}(x_0 + \theta_1 h_1, y_0 + h_2) - \frac{\partial f}{\partial x}(x_0, y_0)\right] \frac{h_1}{\|(h_1, h_2)\|} \\ &+ \left[\frac{\partial f}{\partial y}(x_0, y_0 + \theta_2 h_2) - \frac{\partial f}{\partial y}(x_0, y_0)\right] \frac{h_2}{\|(h_1, h_2)\|} \end{aligned}$$

Observemos que

$$\frac{|h_i|}{\|(h_1, h_2)\|} \leq 1 \quad \forall (h_1, h_2) \neq (0, 0), i = 1, 2.$$

y que, debido a la continuidad asumida de las derivadas parciales en (x_0, y_0), cuando $(h_1, h_2) \to (0, 0)$ los dos corchetes de la expresión anterior tienden a cero, obtenemos finalmente que

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = 0$$

lo que nos dice que la función f es diferenciable en (x_0, y_0). Q.E.D.

Ejemplo 7. Las derivadas parciales de la función polinomial $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sum_{i,j} a_{ij} x^i y^j$, también son funciones polinomiales. Como éstas son continuas, entonces la función f es diferenciable, como ya se había dicho en el ejemplo 2. ■

Ejemplo 8. La función $f(x, y) = e^{-(x^2+y^2)}$ tiene por derivadas parciales a

$$\frac{\partial f}{\partial x} = -2xe^{-(x^2+y^2)} \qquad \frac{\partial f}{\partial y} = -2ye^{-(x^2+y^2)}$$

que son funciones continuas. Entonces f es diferenciable (Ver ejemplo 5). ■

Ejemplo 9. La función $f: \mathbb{R}^3 \to \mathbb{R}$ dada por $f(x, y, z) = \cos(x + y^2 + z^3)$ tiene por derivadas parciales a

$$\frac{\partial f}{\partial x} = -\operatorname{sen}(x + y^2 + z^3) \qquad \frac{\partial f}{\partial y} = -2y \operatorname{sen}(x + y^2 + z^3)$$

$$\frac{\partial f}{\partial z} = -3z^2 \operatorname{sen}(x + y^2 + z^3)$$

que son funciones continuas. Entonces f es diferenciable. ■

Ejercicios (Capítulo 2, Sección 6)

Para cada una de las funciones dadas en los ejercicios 1–10, escriba la expresión del residuo de la definición de diferenciabilidad en el punto en cuestión. Pruebe que la función es diferenciable.

1. $f(x, y) = 5$, $\mathbf{p} = (x_0, y_0)$
2. $f(x, y) = 3x$, $\mathbf{p} = (x_0, y_0)$
3. $f(x, y) = 8y$, $\mathbf{p} = (x_0, y_0)$
4. $f(x, y) = 4x - 10y$, $\mathbf{p} = (x_0, y_0)$
5. $f(x, y) = 3x^2 + 9y^2$, $\mathbf{p} = (1, 2)$
6. $f(x, y) = 4x^2y^3$, $\mathbf{p} = (1, 1)$
7. $f(x, y) = x \operatorname{sen} y$, $\mathbf{p} = (0, 0)$
8. $f(x, y, z) = x^3 + 3y^2 + z^2$, $\mathbf{p} = (1, 0, 1)$
9. $f(x, y, z) = xy^2z^3$, $\mathbf{p} = (2, 1, 1)$

10. $f(x, y, z) = e^{x+y+z}$, $\mathbf{p} = (0, 0, 0)$

11. Sean $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones definidas en el conjunto abierto U de $\mathbb{R}^n$, diferenciables en $\mathbf{p} = (x_0, y_0) \in U$.

a. Demuestre que la función suma $f + g: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $(f + g)(x, y) = f(x, y) + g(x, y)$ es una función diferenciable en $\mathbf{p}$.

b. Demuestre que si $g(x_0, y_0) \neq 0$, la función cociente $\frac{f}{g}: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $\left(\frac{f}{g}\right)(x, y) = \frac{f(x, y)}{g(x, y)}$ es diferenciable en $\mathbf{p}$.

En los ejercicios 12–15, utilice los resultados establecidos en el teorema 2.6.2, junto con el hecho de que "composición de funciones diferenciables es diferenciable", para argumentar por qué las funciones dadas son diferenciables (en todo su dominio).

12. $f(x, y) = \ln(1 + x^2 + y^2)$

13. $f(x, y) = 4\operatorname{sen}^2(x + y)\cos^2(x - y)$

14. $f(x, y) = \dfrac{1 + \tanh x}{\cosh y}$

15. $f(x, y) = \arctan \dfrac{1 + x + y}{1 - x - y}$

Sea $g: \mathbb{R} \to \mathbb{R}$ una función diferenciable definida en $\mathbb{R}$. Para cada una de las funciones $z = f(x, y)$ dadas en los ejercicios 16–20, diga dónde están definidas y si son diferenciables (en su dominio).

16. $f(x, y) = 4g(x) + 7g(y)$

17. $f(x, y) = xyg(x)g(y)$

18. $f(x, y) = g(x)g(g(y)) + g(y)g(g(x))$

19. $f(x, y) = \dfrac{g(y)}{1 + g^2(x)}$

20. $f(x, y) = g(xg(y)) + xg(y)g(g(g(x)))$

21. Concluya de nuevo la diferenciabilidad de las funciones de los ejercicios 12–15, calculando sus derivadas parciales y aplicando el teorema 2.6.3.

22. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = |x| + |y|$. ¿Qué aspecto tiene la gráfica de f? Demuestre que esta función no es diferenciable en el origen. ¿En qué otros puntos no es diferenciable?

23. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$ y sea $\mathbf{p} = (x_0, y_0) \in U$. Suponga que el límite $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ no existe. Demuestre que f no es diferenciable en $\mathbf{p}$.

24. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$ y sea $\mathbf{p} = (x_0, y_0)$ un punto frontera de U. Demuestre que f no es diferenciable en $\mathbf{p}$.

25. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$. Demuestre que f es continua en cualquier punto $\mathbf{p} \in U$, en la dirección de cualquier vector $\mathbf{v} \in \mathbb{R}^2$ (ver ejercicio 81 de la sección 3).

26. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$ y sea $\mathbf{p} = (x_0, y_0)$ un punto de U. Suponga que hay un vector $\mathbf{v} = (a, b) \in \mathbb{R}^2$ no nulo para el que la derivada direccional de f en $\mathbf{p}$, en la dirección de $\mathbf{v}$, no existe. Demuestre que f no es diferenciable en $\mathbf{p}$.

27. (A manera de recapitulación: ¿qué implica qué?). Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$, y sea $\mathbf{p}$ un punto de U. A continuación se dan 8 afirmaciones sobre la función f.

a. f es diferenciable en $\mathbf{p}$.

b. f es continua respecto de su primera variable en $\mathbf{p}$.

c. f es continua respecto de su segunda variable en $\mathbf{p}$.

d. f es continua en $\mathbf{p}$ en la dirección de algún vector $\mathbf{v} \in \mathbb{R}^2$.

e. f es continua en $\mathbf{p}$ en la dirección de todo vector $\mathbf{v} \in \mathbb{R}^2$.

f. f tiene derivadas parciales en $\mathbf{p}$.

g. f tiene derivadas direccionales en $\mathbf{p}$ en la dirección de cualquier vector $\mathbf{v} \in \mathbb{R}^2$.

h. f tiene derivadas parciales continuas en alguna bola B contenida en U con centro en $\mathbf{p}$.

Llene el siguiente cuadro, indicando con una V en la línea i y columna j, cuando la afirmación de la línea i implique la afirmación de la columna j, y con una F cuando no la implique. Por ejemplo, la afirmación **a** implica la afirmación **f**, pero la afirmación **f** no implica la **a**. Estas respuestas ya aparecen en la tabla.

	a	b	c	d	e	f	g	h
a						V		
b								
c								
d								
e								
f	F							
g								
h								

28. En este ejercicio se muestra que la afirmación recíproca del teorema 2.6.3 es falsa. Es decir, el hecho de que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ sea diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in U$, no implica que las derivadas parciales de f sean continuas en $\mathbf{p}$. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} (x^2 + y^2) \operatorname{sen} \dfrac{1}{\sqrt{x^2 + y^2}} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

a. Demuestre que las derivadas parciales de esta función están dadas por $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}: \mathbb{R}^2 \to \mathbb{R}$

$$\frac{\partial f}{\partial x} = \begin{cases} 2x \operatorname{sen} \dfrac{1}{\sqrt{x^2 + y^2}} - \dfrac{x}{\sqrt{x^2 + y^2}} \cos \dfrac{1}{\sqrt{x^2 + y^2}} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} 2y \operatorname{sen} \dfrac{1}{\sqrt{x^2 + y^2}} - \dfrac{y}{\sqrt{x^2 + y^2}} \cos \dfrac{1}{\sqrt{x^2 + y^2}} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

b. Demuestre que las derivadas parciales de f son *discontinuas* en el origen, probando que el límite de ellas cuando (x, y) tiende a $(0, 0)$ no existe.

c. Constate que el residuo de la defínción de diferenciabilidad aplicada a f en el origen se ve como

$$r(h_1, h_2) = (h_1^2 + h_2^2) \operatorname{sen} \frac{1}{\sqrt{h_1^2 + h_2^2}}$$

d. Demuestre que

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = 0$$

y concluya entonces que la función es diferenciable en el origen.

29. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$, y sea $\mathbf{x}_0 \in U$. Considere la función ρ: $B \to \mathbb{R}$ definida en una bola de $\mathbb{R}^n$ con centro en $\mathbf{0}$, dada por

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)h_i + \rho(\mathbf{h})\|\mathbf{h}\|$$

Demuestre que la función f es diferenciable en $\mathbf{x}_0$ si y sólo si la función ρ es tal que $\rho(\mathbf{0}) = 0$, y $\lim_{h\to 0} \rho(\mathbf{h}) = 0$.

30. Determine la función $\rho(\mathbf{h})$ del ejercicio anterior para cada una de las funciones de los ejercicios 1–10. Verifique nuevamente que las funciones dadas en estos ejercicios son diferenciables (en los puntos indicados).

31. (Un breve curso "hágalo usted mismo" sobre funciones de variable compleja. parte (I): diferenciabilidad). En este ejercicio trabajaremos con el conjunto de números complejos $\mathbb{C} = \{x + iy | x, y \in \mathbb{R},\ i^2 = -1\}$. Recordemos primeramente algunos hechos básicos sobre este conjunto. Dado el número complejo $z = x + iy$, decimos que x es su *parte real*, y que y es su *parte imaginaria*. Se escribe $x = \operatorname{Re} z$, $y = \operatorname{Im} z$. Dos números complejos z_1 y z_2 se dicen ser *iguales*, lo cual se escribe como $z_1 = z_2$, si tienen la misma parte real y la misma parte imaginaria. Se define el *módulo* del número complejo $z = x + iy$, denotado por $|z|$, como $|z| = \sqrt{x^2 + y^2}$. Dado el complejo $z = x + iy$, se define su *conjugado*, denotado por $\bar{z}$, como $\bar{z} = x - iy$. En el conjunto $\mathbb{C}$ se introducen las operaciones de *suma* y *producto* de la siguiente manera: si $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, entonces $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$, $z_1z_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$. Se demuestra que $\mathbb{C}$ con estas operaciones forma un campo. Esto se comprobará en el primer inciso de este ejercicio, en el cual, además, se recuerdan algunos resultados elementales sobre el manejo algebraico con números complejos.

a. Sean z, z_1, z_2, z_3 números complejos, y a, b, c números reales. En el inciso **a8** se probará que todo $z \in \mathbb{C}$, $z \neq 0$, tiene un inverso multiplicativo z^{-1}, tal que $zz^{-1} = 1$. Defina el *cociente* del complejo z_1 entre el complejo z_2 (no nulo), denotado por $\frac{z_1}{z_2}$ o z_1/z_2, como $z_1z_2^{-1}$. Demuestre que:

a1. $z_1 + z_2 = z_2 + z_1$

a2. $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$

a3. $z + 0 = 0$, en donde $0 = 0 + i0$

a4. dado $z \in \mathbb{C}$, existe $-z \in \mathbb{C}$ tal que $z + (-z) = 0$

a5. $z_1 z_2 = z_2 z_1$

a6. $z_1(z_2 z_3) = (z_1 z_2) z_3$

a7. $1z = z$, en donde $1 = 1 + i0$

a8. dado $z \in \mathbb{C}$, $z \neq 0$, existe z^{-1} tal que $z^{-1}z = 1$

a9. $z(z_1 + z_2) = zz_1 + zz_2$

a10. $|z| = |\bar{z}|$

a11. $z\bar{z} = |z|^2$

a12. $|z\bar{z}| = |z||\bar{z}|$

a13. $|z_1/z_2| = |z_1|/|z_2|$

a14. $|z_1 + z_2| \leq |z_1| + |z_2|$

b. Identifique el número complejo $z = x + iy$, como el punto del plano $\mathbb{R}^2$ cuyas coordenadas son (x, y). Pensando en esta identificación es que hablaremos del "plano complejo". Describa los siguientes conjuntos de puntos en el plano complejo:

b1. $\{z \in \mathbb{C} \mid \operatorname{Re} z = 0\}$

b2. $\{z \in \mathbb{C} \mid \operatorname{Im} z = 1\}$

b3. $\{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$

b4. $\{z \in \mathbb{C} \mid \operatorname{Im} z < \operatorname{Re} z\}$

b5. $\{z \in \mathbb{C} \mid |z| = 1\}$

b6. $\{z \in \mathbb{C} \mid |z| \leq 1\}$

b7. $\{z \in \mathbb{C} \mid |z| \geq 1\}$

c. Sea z_0 un complejo dado. Defina la *bola abierta con centro* en z_0 y *radio* $r > 0$, denotada por $B(z_0, r)$ como el conjunto $\{z \in \mathbb{C} \mid |z - z_0| < r\}$. Identifique geométricamente este conjunto en el plano complejo.

d. Un subconjunto U de $\mathbb{C}$ se dice ser *abierto*, si para cada $z \in U$ existe una bola $B(z, r)$ tal que $B(z, r) \subset U$. De los conjuntos dados en el inciso b, diga cuáles son abiertos.

e. Una función del tipo $f: U \subseteq \mathbb{C} \to \mathbb{C}$, que a cada número complejo z de su dominio $U \subseteq \mathbb{C}$ le asocia un número complejo $f(z) \in \mathbb{C}$ bien determinado, se dice ser una *función (compleja) de (una) variable compleja*. Demuestre que existen funciones reales $u, v: U \subseteq \mathbb{C} \to \mathbb{R}$, tales que $f(z) = u(z) + iv(z)$. Si consideramos el dominio de estas funciones como parte del plano $\mathbb{R}^2$, podemos decir que las funciones $u, v: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son tales que $f(x + iy) = u(x, y) + iv(x, y)$. Nótese entonces que u y v son del tipo de funciones que hemos estudiado en este capítulo. A u se le llama la *parte real de* f, y a v se le llama la *parte imaginaria de* f. Identifique las funciones u y v para cada una de las funciones de variable compleja dadas a continuación:

e1. $f(z) = z$

e2. $f(z) = z^2$

e3. $f(z) = \dfrac{1}{z}$, $z \neq 0$

e4. $f(z) = (3z + 2)^2$

e5. $f(z) = \bar{z}$

e6. $f(z) = |z|$

f. Sea $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$, $f = u + iv$, una función definida en el conjunto abierto U de $\mathbb{C}$. Sea z_0 un punto de U. Se dice que el límite de f, cuando z tiende a z_0, es $L(\in \mathbb{C})$, lo cual se escribe como $\lim_{z \to z_0} f(z) = L$, si dado $\epsilon > 0$, existe $\delta > 0$, tal que si $0 < |z - z_0| < \delta$, entonces $|f(z) - L| < \epsilon$. Demuestre que si f es una función polinomial, es decir, es una función del tipo $f(z) = a_n z^n + \cdots + a_1 z + a_0$, entonces $\lim_{z \to z_0} f(z) = f(z_0)$. Más aún, demuestre que este resultado es cierto para funciones racionales, es decir, funciones que son cocientes de dos funciones polinomiales, con z_0 en el dominio de la función. Calcule:

f1. $\displaystyle\lim_{z \to (1+i)} z^2$

f2. $\displaystyle\lim_{z \to 1} \frac{z^2+3}{z-2i}$

f3. $\displaystyle\lim_{z \to (1+i)} \frac{z^3 - 3z + 2}{z^2 + 4i}$

g. Sea $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ una función definida en el conjunto abierto U de $\mathbb{C}$. Sea $z_0 \in U$. Se define la *derivada* de la función f en z_0, denotada por $f'(z_0)$, como el límite

$$f'(z_0) = \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

cuando se da este límite. Escriba $f(z) = u(x, y) + iv(x, y)$, $z_0 = (x_0, y_0) = x_0 + iy_0$, $\mathbf{h} = (h_1, h_2) = h_1 + ih_2$. Demuestre que $f'(z_0)$ se puede escribir como $a + ib$, donde

$$a = \lim_{h \to 0} \frac{u(z_0 + h) - u(z_0)}{h}, \qquad b = \lim_{h \to 0} \frac{v(z_0 + h) - v(z_0)}{h}$$

h. Considere el primero de los límites del inciso anterior

$$a = \lim_{h \to 0} \frac{u(z_0 + h) - u(z_0)}{h}$$

Ponga $\mathbf{h} = (h_1, 0) = h_1 \in \mathbb{R}$. Demuestre entonces que

$$a = \lim_{h_1 \to 0} \frac{u(x_0 + h_1, y_0) - u(x_0, y_0)}{h_1} = \frac{\partial u}{\partial x}(x_0, y_0)$$

Ponga $\mathbf{h} = (0, h_2) = ih_2$. Demuestre

$$a = \lim_{h_2 \to 0} \frac{u(x_0, y_0 + h_2) - u(x_0, y_0)}{ih_2} = \frac{1}{i}\frac{\partial u}{\partial y}(x_0, y_0) = -i\frac{\partial u}{\partial y}(x_0, y_0)$$

i. Considere el segundo de los límites del inciso **g**

$$b = \lim_{h \to 0} \frac{v(z_0 + h) - v(z_0)}{h}$$

Ponga $\mathbf{h} = (h_1, 0) = h_1 \in \mathbb{R}$. Demuestre entonces que

$$b = \lim_{h_1 \to 0} \frac{v(x_0 + h_1, y_0) - v(x_0, y_0)}{h_1} = \frac{\partial v}{\partial x}(x_0, y_0)$$

Ponga $\mathbf{h} = (0, h_2) = ih_2$. Demuestre

$$b = \lim_{h\to 0} \frac{v(x_0, y_0 + h_2) - v(x_0, y_0)}{ih_2} = \frac{1}{i}\frac{\partial v}{\partial y}(x_0, y_0) = -i\frac{\partial v}{\partial y}(x_0, y_0)$$

j. Use los resultados de los dos incisos anteriores para demostrar que $f'(z_0)$ se puede escribir como

$$\frac{\partial u}{\partial x}(x_0, y_0) + i\frac{\partial v}{\partial x}(x_0, y_0)$$

y como

$$-i\frac{\partial u}{\partial y}(x_0, y_0) + i\left(-i\frac{\partial v}{\partial y}(x_0, y_0)\right) = \frac{\partial v}{\partial y}(x_0, y_0) - i\frac{\partial u}{\partial y}(x_0, y_0)$$

Concluya entonces que si $f'(z_0)$ existe, se deben cumplir las igualdades

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0)$$
$$\frac{\partial v}{\partial x}(x_0, y_0) = -\frac{\partial u}{\partial y}(x_0, y_0)$$

Estas se llaman *Ecuaciones de Cauchy-Riemann.* En el siguiente inciso se establecen de nuevo estas ecuaciones siguiendo la idea de diferenciabilidad de funciones estudiada en esta sección.

k. Reconsiderando la definición establecida en el inciso g, diremos que la función $f: U \subseteq \mathbb{C} \to \mathbb{C}$, definida en el conjunto abierto U de $\mathbb{C}$, es *diferenciable* en $z_0 \in U$, si existe un número complejo $f'(z_0) = a + ib$, tal que definiendo $r(h)$ como

$$f(z_0 + h) = f(z_0) + f'(z_0)h + r(h)$$

se tiene

$$\lim_{h\to 0} \frac{r(h)}{h} = 0$$

Poniendo $f(z) = u(x, y) + iv(x, y)$, $z_0 = (x_0, y_0) = x_0 + iy_0$, $h = h_1 + ih_2$, $r(h) = r_1(h) + ir_2(h)$, demuestre que podemos separar las partes real e imaginaria de la expresión anterior y escribir entonces que

$$u(z_0 + h) = u(z_0) + ah_1 - bh_2 + r_1(h), \qquad \text{donde } \lim_{h\to 0} \frac{r_1(h)}{h} = 0$$
$$v(z_0 + h) = v(z_0) + bh_1 + ah_2 + r_2(h), \qquad \text{donde } \lim_{h\to 0} \frac{r_2(h)}{h} = 0$$

Concluya entonces que si la función compleja f es diferenciable en z_0, las partes real $u(x, y)$ e imaginaria $v(x, y)$ de f son funciones diferenciables (según se estudió en esta sección). Comparando las dos expresiones anteriores (los coeficientes a y b que multiplican a las coordenadas de h), concluya que se debe tener

$$a = \frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0)$$
$$b = \frac{\partial v}{\partial x}(x_0, y_0) = -\frac{\partial u}{\partial y}(x_0, y_0)$$

Estas son las ecuaciones de Cauchy-Riemann establecidas en el inciso anterior.

l. Si la función $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ es diferenciable para todo $z \in U$, se dice que es *diferenciable* (en U), o bien, que es una *función holomorfa*. El cumplimiento de las ecuaciones de Cauchy-Riemann en un punto z_0 son entonces una condición *necesaria* para que la función f sea diferenciable en z_0. El siguiente ejemplo muestra que el cumplimiento de estas ecuaciones no es suficiente para garantizar la diferenciabilidad de la función: sea $f\colon \mathbb{C} \to \mathbb{C}$ la función definida como $f(z) = z^5/|z|^4$, si $z \neq 0$, $f(0) = 0$. Demuestre que esta función satisface las ecuaciones de Cauchy-Riemann en $z = 0$, pero que no es diferenciable en este punto (es decir, $f'(0)$ no existe).

m. Aplique las ecuaciones de Cauchy-Riemann para demostrar que la función $f(z) = \bar{z}$ no es diferenciable en punto alguno.

n. Demuestre que la función $f(z) = z \operatorname{Re} z$ es diferenciable solamente en el punto $z = 0$.

o. Verifique que las siguientes funciones satisfacen las ecuaciones de Cauchy-Riemann en todo punto de su dominio:

o1. $f(z) = z^2$

o2. $f(z) = z^3$

o3. $f(z) = \dfrac{1}{z}$

o4. $f(z) = (4z + i)^2$

p. Se puede demostrar que si la función $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ es tal que sus partes real e imaginaria u, $v\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son funciones de clase $\mathscr{C}^1$ y satisfacen las ecuaciones de Cauchy-Riemann en un punto $z_0 \in U$ y en *una bola abierta con centro* en z_0 *contenida* en U, entonces la función f es diferenciable en z_0. Use este hecho para concluír que las funciones del inciso anterior son diferenciables en todo su dominio (es decir, son funciones holomorfas).

(Este ejercicio continuará en el ejercicio 34 de la sección 12 de este capítulo).

2.7 Diferenciabilidad y derivadas direccionales

Retomemos la discusión sobre derivadas direccionales que se presentó en la sección 2.5. En esa sección se calcularon las derivadas direccionales de las funciones utilizando solamente la definición ahí presentada de ese concepto. A partir de esta sección las funciones con las que vamos a trabajar son (como ya dijimos) diferenciables —a menos que se indique lo contrario—. En tal caso, el cálculo de derivadas direccionales es mucho más amable, sin hacer uso de la definición.

Tomemos la función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable, definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $\mathbf{x}_0 \in U$ y sea $\mathbf{v} \in \mathbb{R}^n$ el vector (unitario) en cuya dirección queremos calcular la derivada de la función f en el punto $\mathbf{x}_0$. Siendo f diferenciable en $\mathbf{x}_0$ se tiene que el residuo $r(\mathbf{h})$ definido por la expresión

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i + r(\mathbf{h})$$

tiene la propiedad de que

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Escribamos el vector $\mathbf{h} = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ como $\mathbf{h} = t\mathbf{v}$, $t \in \mathbb{R}$ (con t suficientemente pequeño para que $t\mathbf{v} \in U$). Entonces $h_i = tv_i$ donde $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. Nos queda entonces que

$$f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)(tv_i) + r(t\mathbf{v})$$

de donde

$$\frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)v_i + \frac{r(t\mathbf{v})}{t}$$

Nótese que decir $\mathbf{h} \to \mathbf{0}$ equivale a pedir que $t \to 0$. Además $\|\mathbf{h}\| = \|t\mathbf{v}\| = |t|\|\mathbf{v}\| = |t|$ pues el vector $\mathbf{v}$ es un vector unitario. Entonces si se toma límite cuando $\mathbf{h} \to \mathbf{0}$ (es decir, cuando $t \to 0$) en la última expresión, nos queda (teniendo en cuenta que los primeros sumandos del 2° miembro no dependen de $\mathbf{h}$ ni de t)

$$\lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)v_i \pm \lim_{h\to 0} \frac{r(\mathbf{h})}{\|\mathbf{h}\|}$$

El límite que aparece en el 2° miembro es igual a cero, pues la función f es diferenciable. En el primer miembro aparece la definición de la derivada de la función f en el punto $\mathbf{x}_0$ en la dirección del vector $\mathbf{v}$. Tenemos pues que

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)v_i \qquad \text{(D)}$$

Y así, con esta sencilla fórmula, podemos calcular las derivadas direccionales de la función f.

Ejemplo 1. Retomemos el ejemplo 1 de la sección 5. Se quiere calcular la derivada de la función $f: \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = x^2 + y^2$ en un punto arbitrario $(x, y) \in \mathbb{R}^2$ en la dirección del vector $\mathbf{v} = (\cos\theta, \operatorname{sen}\theta)$. Según la fórmula (D) tenemos

$$\frac{\partial f}{\partial \mathbf{v}} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\operatorname{sen}\theta = 2x\cos\theta + 2y\operatorname{sen}\theta$$

De igual forma, en el ejemplo 3 de la sección 5 se consideró la función $f: \mathbb{R}^3 \to \mathbb{R}$, dada por $f(x, y, z) = 2x^3 + 7y^2 + 9z^2$ y se quería calcular su derivada direccional en un punto arbitrario $(x, y, z) \in \mathbb{R}^3$ en la dirección del vector unitario $\mathbf{v} = (a, b, c) \in \mathbb{R}^3$. Se tiene entonces

$$\frac{\partial f}{\partial \mathbf{v}} = \frac{\partial f}{\partial x}a + \frac{\partial f}{\partial y}b + \frac{\partial f}{\partial z}c = 6x^2a + 14yb + 18zc$$

tal como se obtuvo en aquélla ocasión. ■

Ejemplo 2. Se quiere calcular la derivada de la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \operatorname{sen}(x^2 + y^2)$ en el punto $\mathbf{p} = (1, 1)$ en la dirección que va de éste punto al punto $\mathbf{q} = (3, 2)$. Según la fórmula (D) todo lo que necesitamos son las derivadas parciales de la función f calculadas en $\mathbf{p}$ y las coordenadas del vector unitario $\mathbf{v}$ que marca la dirección en que se quiere calcular la derivada. El

vector $\mathbf{u} = \mathbf{q} - \mathbf{p} = (3, 2) - (1, 1) = (2, 1)$ está en la dirección deseada, de modo que el vector $\mathbf{v}$ será (normalizando al vector $\mathbf{u}$)

$$\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{1}{\sqrt{5}}(2, 1) = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right) = (v_1, v_2)$$

Por otra parte, las derivadas parciales de la función son

$$\frac{\partial f}{\partial x} = 2x\cos(x^2 + y^2), \qquad \frac{\partial f}{\partial y} = 2y\cos(x^2 + y^2)$$

las cuales, evaluadas en $\mathbf{p}$ son

$$\frac{\partial f}{\partial x}(1, 1) = 2\cos 2, \qquad \frac{\partial f}{\partial y}(1, 1) = 2\cos 2$$

Entonces

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(1, 1) &= \frac{\partial f}{\partial x}(1, 1)v_1 + \frac{\partial f}{\partial y}(1, 1)v_2 \\ &= (2\cos 2)\left(\frac{2}{\sqrt{5}}\right) + (2\cos 2)\left(\frac{1}{\sqrt{5}}\right) = \frac{6}{\sqrt{5}}\cos 2\end{aligned}$$ ■

Ejemplo 3. Consideremos ahora la función $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2$, y fijemos nuestra atención en el punto $(1, 1, 2)$ de la superficie que representa (un paraboloide con vértice en el origen que abre hacia arriba). La curva de nivel de $z = x^2 + y^2$ que "pasa por $\mathbf{p}$" (corresponde al nivel 2) es la circunferencia $x^2 + y^2 = 2$. Queremos calcular la derivada de la función f en el punto $\mathbf{p} = (1, 1)$ —el cual pertenece a la curva del nivel 2— en la dirección de la tangente a la curva $x^2 + y^2 = 2$ en ese punto. Es decir, queremos calcular la derivada de la función f en el punto $\mathbf{p}$ en la dirección de la curva de nivel que contiene a $\mathbf{p}$.

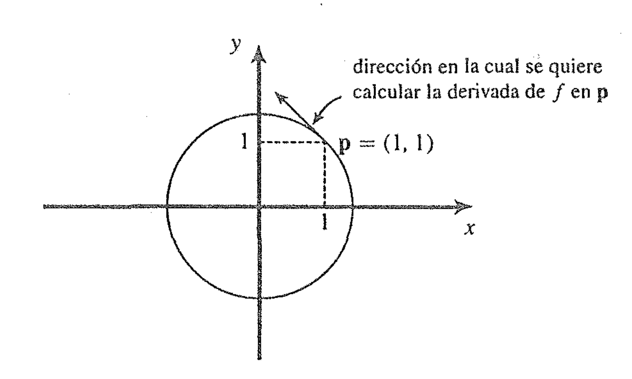


Figura 1. Gráfica de la tangente a la curva de nivel en el punto $\mathbf{p}$.

Un vector unitario en la dirección requerida es $\mathbf{v} = (-\sqrt{2}/2, \sqrt{2}/2)$. (¿Cómo se obtiene? Fácilmente: la dirección buscada debe ser ortogonal al punto $\mathbf{p}$ del círculo $x^2 + y^2 = 2$; un vector ortogonal a $(1, 1)$ es cualquier vector (x, y) que satisfaga que su producto punto con $(1, 1)$ sea cero —condición de ortogonalidad—, i.e., $x + y = 0$; por ejemplo el $(-1, 1)$. Sólo falta entonces hacerlo unitario, y así se obtiene el vector $\mathbf{v} = \frac{1}{\|(-1,1)\|}(-1, 1) = (-\sqrt{2}/2, \sqrt{2}/2)$.

Por otra parte, las derivadas parciales de la función f evaluadas en $\mathbf{p}$ son:

$$\frac{\partial f}{\partial x}(1, 1) = 2, \qquad \frac{\partial f}{\partial y}(1, 1) = 2$$

de modo que la derivada de f en la dirección de $\mathbf{v}$ es

$$\frac{\partial f}{\partial \mathbf{v}}(1, 1) = \frac{\partial f}{\partial x}(1, 1)v_1 + \frac{\partial f}{\partial y}(1, 1)v_2 = 2\left(-\frac{\sqrt{2}}{2}\right) + 2\left(\frac{\sqrt{2}}{2}\right) = 0$$

Recapacitemos un poco en el resultado obtenido; si estamos en un punto de alguna montaña (digamos que nuestra superficie $z = x^2 + y^2$ "volteada") y caminamos sobre ella en una línea de nivel —levantada al punto en cuestión— por definición de curva de nivel, los valores de la función no cambian durante el recorrido. Es decir, yendo por una curva de nivel, los valores de la función se mantienen constantes, y, por tanto, se espera que la derivada de la función en esa dirección —que mide la velocidad de la variación de la función en ese punto en dicha dirección— sea cero, como ocurrió en este ejemplo. Este es un hecho general que se demostrará posteriormente.

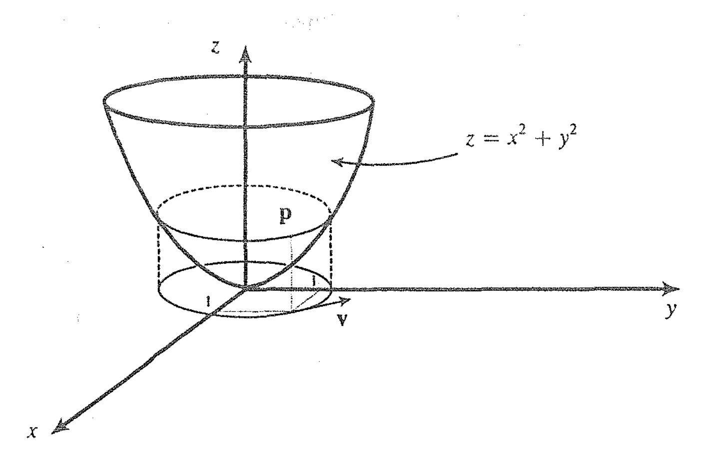


Figura 2. Derivada direccional sobre la curva de nivel de un paraboloide.

Generalizando este ejemplo, si $\mathbf{p} = (x_0, y_0, c)$ es un punto cualquiera de la superficie $z = x^2 + y^2$ (y entonces $x_0^2 + y_0^2 = c$) por el cual pasa la curva de nivel $x^2 + y^2 = c$ (levantada a $\mathbf{p}$), un vector unitario en la dirección de la tangente a la curva en $\mathbf{p}$ es $\mathbf{v} = (-\operatorname{sen}\theta_0, \cos\theta_0)$ en donde $\tan\theta_0 = \frac{y_0}{x_0}$, de modo que la derivada de la función f en $\mathbf{p}$ en la dirección de $\mathbf{v}$ es

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(x_0, y_0) &= \frac{\partial f}{\partial x}(x_0, y_0)v_1 + \frac{\partial f}{\partial y}(x_0, y_0)v_2 \\ &= 2x_0(-\operatorname{sen}\theta_0) + 2y_0(\cos\theta_0) \\ &= 2x_0\left(-\frac{y_0}{\sqrt{c}}\right) + 2x_0\left(\frac{y_0}{\sqrt{c}}\right) = 0\end{aligned}$$

tal como dijimos anteriormente que tenía que ocurrir. ■

Ejemplo 4. Consideremos la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = e^{-(x^2+y^2)}$. Sea $\mathbf{v} = (v_1, v_2)$ un vector unitario dado en el espacio $\mathbb{R}^2$. Calculemos la derivada de la función f en el origen en la dirección del vector $\mathbf{v}$, las derivadas parciales son

$$\frac{\partial f}{\partial x} = -2xe^{-(x^2+y^2)}, \qquad \frac{\partial f}{\partial y} = -2ye^{-(x^2+y^2)}$$

de modo que en el origen ambas derivadas son nulas.

Entonces

$$\frac{\partial f}{\partial \mathbf{v}}(0, 0) = 0v_1 + 0v_2 = 0$$

Es decir, la derivada de la función f en el origen, *en cualquier dirección* (el vector $\mathbf{v}$ fue arbitrario) es cero. ■

La fórmula 7.1 se puede leer de manera vectorial si la identificamos como un producto punto de dos vectores: uno de ellos el vector unitario $\mathbf{v}$ que marca la dirección en la cual calcularemos la derivada de la función y el otro, el vector de $\mathbb{R}^n$ cuya i-ésima coordenada es la derivada parcial de la función f respecto de su i-ésima variable, evaluada en el punto en cuestión, en el que se está calculando la derivada direccional. Este último vector jugará, a continuación, un papel fundamental. El objetivo de la próxima sección es presentarlo de manera formal y estudiar algunas de sus propiedades.

Apéndice El teorema de Euler sobre funciones homogéneas

En este apéndice vamos a estudiar un teorema que establece una relación sobre cierto tipo de funciones llamadas homogéneas, cuya demostración tiene que ver con las ideas manejadas en esta sección. Como siempre, nuestras funciones serán diferenciables.

Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Se dice que esta función es *homogénea de grado* α (α como número real) si se cumple que

$$f(t\mathbf{x}) = t^\alpha f(\mathbf{x}), \ \forall \mathbf{x} \in U, \ \forall t > 0$$

Veamos algunos ejemplos.

Ejemplo 1. La función $f: \mathbb{R} \to \mathbb{R}$, dada por $f(x) = 3x$ es homogénea de grado 1 pues

$$f(tx) = 3(tx) = t(3x) = tf(x), \forall x \in \mathbb{R}, \forall t > 0$$

Del mismo modo la función $g: \mathbb{R} \to \mathbb{R}$, dada por $g(x) = 10x^2$ es homogénea de grado 2, y la función $h: \mathbb{R} \to \mathbb{R}$, dada por $h(x) = 8x + 1$, no es homogénea. ■

Ejemplo 2. La función $f: \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = 8x^3 + 3x^2y + 5y^3$ es homogénea de grado 3 pues

$$\begin{aligned} f(tx, ty) &= 8(tx)^3 + 3(tx)^2(ty) + 5(ty)^3 \\ &= t^3(8x^3 + 3x^2y + 5y^3) = t^3 f(x, y), \ \forall (x, y) \in \mathbb{R}^2, \ \forall t > 0 \end{aligned}$$

La función $g: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}$, dada por

$$g(x, y) = \operatorname{sen} \frac{x^2 - y^2}{x^2 + y^2}$$

es homogénea de grado 0, pues

$$\begin{aligned} g(tx, ty) &= \operatorname{sen} \frac{(tx)^2 - (ty)^2}{(tx)^2 + (ty)^2} = \operatorname{sen} \frac{t^2(x^2 - y^2)}{t^2(x^2 + y^2)} \\ &= \operatorname{sen} \frac{x^2 - y^2}{x^2 + y^2} = t^0 f(x, y), \ \forall (x, y) \in \mathbb{R}^2, \ \forall t > 0. \end{aligned}$$

La función $h: \mathbb{R}^2 \to \mathbb{R}$, dada por $h(x, y) = 8xy + 10x^3 + 1$ *no* es homogénea. ■

Ejemplo 3. La función $f: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$f(x, y, z) = x^3y + 8xyz^2 + 10y^3z$$

es homogénea de grado 4, en tanto que la función $g: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$g(x, y, z) = \sqrt[3]{x^5 + 8x^4y + 15xyz^3 + z^5}$$

es homogénea de grado 5/3. Las verificaciones son análogas a los ejemplos anteriores y quedan como ejercicio para el lector. ■

El resultado que queremos establecer en este apéndice es el siguiente.

Teorema (de Euler sobre funciones homogéneas). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función homogénea de grado α definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $\mathbf{x}_0$ un punto no nulo de U, y sea $\mathbf{v} = \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|}$ el vector unitario en la dirección de $\mathbf{x}_0$. Entonces (si la derivada direccional $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ existe)

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \frac{\alpha}{\|\mathbf{x}_0\|} f(\mathbf{x}_0)$$

Demostración. Directamente de la definición de derivada direccional:

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) &= \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} \\ &= \lim_{t \to 0} \frac{f\left(\mathbf{x}_0 + t\frac{\mathbf{x}_0}{\|\mathbf{x}_0\|}\right) - f(\mathbf{x}_0)}{t} \\ &= \lim_{t \to 0} \frac{f\left(\frac{\|\mathbf{x}_0\| + t}{\|\mathbf{x}_0\|}\mathbf{x}_0\right) - f(\mathbf{x}_0)}{t} \end{aligned}$$

Siendo f homogénea de grado α tenemos que

$$f\left(\frac{\|\mathbf{x}_0\|+t}{\|\mathbf{x}_0\|}\mathbf{x}_0\right) = \left(\frac{\|\mathbf{x}_0\|+t}{\|\mathbf{x}_0\|}\mathbf{x}_0\right)^{\alpha} f(\mathbf{x}_0)$$

de modo que

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) &= \lim_{t\to 0}\frac{\left(\frac{\|\mathbf{x}_0\|+t}{\|\mathbf{x}_0\|}\mathbf{x}_0\right)^{\alpha} f(\mathbf{x}_0) - f(\mathbf{x}_0)}{t}\\ &= f(\mathbf{x}_0)\lim_{t\to 0}\frac{\left(\frac{\|\mathbf{x}_0\|+t}{\|\mathbf{x}_0\|}\mathbf{x}_0\right)^{\alpha} - 1}{t}\end{aligned}$$

Considere ahora la función $\varphi(x) = x^{\alpha}$. Observe que

$$\begin{aligned}\varphi'(1) &= \lim_{h\to 0}\frac{\varphi(1+h)-\varphi(1)}{h} = \lim_{h\to 0}\frac{(1+h)^{\alpha}-1}{h}\\ &\overset{h=\frac{t}{\|\mathbf{x}_0\|}}{=} \lim_{t\to 0}\frac{\left(1+\frac{t}{\|\mathbf{x}_0\|}\right)^{\alpha}-1}{\frac{t}{\|\mathbf{x}_0\|}}\\ &= \|\mathbf{x}_0\|\lim_{t\to 0}\frac{\left(\frac{\|\mathbf{x}_0\|+t}{\|\mathbf{x}_0\|}\right)^{\alpha}-1}{t}\end{aligned}$$

de modo que

$$\lim_{t\to 0}\frac{\left(\frac{\|\mathbf{x}_0\|+t}{\|\mathbf{x}_0\|}\right)^{\alpha}-1}{t} = \frac{1}{\|\mathbf{x}_0\|}\varphi'(1) = \frac{\alpha}{\|\mathbf{x}_0\|}$$

y así

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \frac{\alpha}{\|\mathbf{x}_0\|}f(\mathbf{x}_0)$$

lo que queríamos probar. Q.E.D.

Corolario Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable, homogénea de grado α, definida en el conjunto abierto U de $\mathbb{R}^n$. Entonces

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n} = \alpha f$$

(las derivadas parciales y la función están evaluadas en un punto arbitrario $\mathbf{x} \in U$).

Demostración. Si $\mathbf{x} \in U$, el vector $\mathbf{v}$ del teorema anterior lo podemos poner como $\mathbf{v} = \frac{1}{\|\mathbf{x}\|}\mathbf{x} = \left(\frac{x_1}{\|\mathbf{x}\|}, \frac{x_2}{\|\mathbf{x}\|}, \ldots, \frac{x_n}{\|\mathbf{x}\|}\right)$, de modo que, según la fórmula (D) se tiene

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) &= \frac{x_1}{\|\mathbf{x}\|}\frac{\partial f}{\partial x_1} + \frac{x_2}{\|\mathbf{x}\|}\frac{\partial f}{\partial x_2} + \cdots + \frac{x_n}{\|\mathbf{x}\|}\frac{\partial f}{\partial x_n} \\ &= \frac{1}{\|\mathbf{x}\|}\left(x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n}\right)\end{aligned}$$

y entonces, por el teorema previamente probado se tiene

$$\frac{1}{\|\mathbf{x}\|}\left(x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n}\right) = \frac{\alpha}{\|\mathbf{x}\|}f$$

de donde

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n} = \alpha f$$

Q.E.D.

Ejemplo 4. Con la función f del ejemplo 2, $f(x, y) = 8x^3 + 3x^2y + 5y^3$ que es homogénea de grado 3 se tiene

$$\begin{aligned}x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} &= x(24x^2 + 6xy) + y(3x^2 + 15y^2) = 24x^3 + 6x^2y + 3x^2y + 15y^3 \\ &= 24x^3 + 9x^2y + 15y^3 = 3(8x^3 + 3x^2y + 5y^3) = 3f(x, y)\end{aligned}$$

tal como aseguraba el corolario anterior. ■

Ejemplo 5. La función $f(x, y) = \frac{x}{y}$ es homogénea de grado 0. Se tiene

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = x\left(\frac{1}{y}\right) + y\left(-\frac{x}{y^2}\right) = \frac{x}{y} - \frac{x}{y} = 0$$

como tenía que ocurrir (según el corolario previamente probado con $\alpha = 0$). ■

Ejercicios (Capítulo 2, Sección 7)

(*) 1 Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en U y sea $\mathbf{u} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ un vector no nulo (no necesariamente de norma 1). Defina la derivada de f en la dirección de $\mathbf{u}$ (en un punto cualquiera $\mathbf{x} \in U$), denotada por $\frac{\partial f}{\partial u}$, como

$$\frac{\partial f}{\partial \mathbf{u}} = \alpha_1\frac{\partial f}{\partial x_1} + \alpha_2\frac{\partial f}{\partial x_2} + \cdots + \alpha_n\frac{\partial f}{\partial x_n}$$

Demuestre que

a. $\dfrac{\partial f}{\partial(c\mathbf{u})} = c\dfrac{\partial f}{\partial \mathbf{u}}$ donde c es un número real no nulo.

b. $\dfrac{\partial f}{\partial(\mathbf{u}_1 + \mathbf{u}_2)} = \dfrac{\partial f}{\partial \mathbf{u}_1} + \dfrac{\partial f}{\partial \mathbf{u}_2}$ donde $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^n$ (ambos no son nulos).

c. Si $\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$ es un vector unitario en la dirección de $\mathbf{u}$, entonces, según el inciso **a** se tiene $\frac{\partial f}{\partial \mathbf{v}} = \frac{1}{\|\mathbf{u}\|}\frac{\partial f}{\partial \mathbf{u}}$. Es decir, la derivada direccional de f en la dirección del vector unitario $\mathbf{v}$ se determina "normalizando" (dividiendo entre la norma) la derivada direccional de f en la dirección del vector $\mathbf{u}$ el *cual tiene la misma dirección que* $\mathbf{v}$. ¿Por qué, si se trata de calcular la derivada *direccional* de f, teniendo $\mathbf{u}$ y $\mathbf{v}$ la misma *dirección*, en el libro se ha definido esta derivada con el vector $\mathbf{v}$ *unitario*?

d. La propiedad $\frac{\partial f}{\partial \mathbf{v}} = \frac{1}{\|\mathbf{u}\|}\frac{\partial f}{\partial \mathbf{u}}$ (que se deduce de lo mostrado en el inciso **a**, ya había sido probada en el ejercicio 16 de la sección 5. Es decir, esta propiedad es válida aún para funciones no diferenciables (con tal de que sus derivadas direccionales existan). ¿Es la propiedad del inciso **b** igualmente válida con sólo suponer que la función posee derivadas direccionales?

En los ejercicios 2–10, calcule la derivada direccional de la función dada en la dirección del vector indicado.

2. $f(x, y) = ax + by$, en un punto $\mathbf{p} = (x_0, y_0)$, en la dirección del vector $\mathbf{u} = (2, 3)$.

3. $f(x, y) = x^2 + y^2$, en el origen, en la dirección del vector $\mathbf{u} = (a, b)$.

4. $f(x, y) = e^{-(x^3+y^3)}$, en el origen, en la dirección del vector $\mathbf{u} = (a, b)$.

5. $f(x, y) = \sqrt{1 - x^2 - y^2}$, en el origen, en la dirección del vector $\mathbf{u} = (a, b)$.

6. $f(x, y) = x \operatorname{sen} y$, en el punto $\mathbf{p} = (1, \pi)$, en la dirección del vector $\mathbf{u} = (-2, 3)$.

7. $f(x, y, z) = x^{y^z}$, en el punto $\mathbf{p} = (1, 1, 1)$, en la dirección del vector $\mathbf{u} = (2, -3, 1)$.

8. $f(x, y, z) = 6 - x^2 - y^2 - z^2$, en el punto $\mathbf{p} = (1, 2, 0)$, en la dirección del vector $\mathbf{u} = (0, 0, 28)$.

9. $f(x, y, z) = ax+by+cz+d$, en el punto $\mathbf{p} = (x_0, y_0, z_0)$, en la dirección del vector $\mathbf{v} = (a, b, c)$.

10. $f(x, y, z) = x \ln y + y \ln z + z \ln x$, en el punto $\mathbf{p} = (1, 1, 1)$, en la dirección del vector $\mathbf{v} = (a, a, a)$ (a un número positivo dado).

11. Calcule la derivada direccional de la función $f(x, y) = 5x^2y^3$ en el punto $\mathbf{p} = (1, 1)$

a. en la dirección del vector que va de $\mathbf{p}$ al punto $(3, -2)$,

b. en la dirección del vector que va de $\mathbf{p}$ al origen,

c. en la dirección del vector tangente al círculo $x^2 + y^2 = 2$ en $\mathbf{p}$,

d. en la dirección del vector $\mathbf{p}$.

12. Calcule la derivada direccional de la función $f(x, y) = x \operatorname{sen} y$ en el punto $(3, 0)$, en la dirección del vector tangente a la parábola $y = x^2$ en el punto $(1, 1)$.

13. Demuestre que la derivada direccional de la función $f(x, y) = \frac{x^2+y^2}{x}$ en los puntos del círculo $x^2 + y^2 - 2y = 0$, en la dirección de la normal a este círculo, es igual a cero.

14. Sea $f(x, y) = x^2 + y^2$. ¿En qué dirección es igual a cero la derivada de esta función en el punto $(1, 1)$?, ¿En qué dirección es igual a cero la derivada de esta función en los puntos del círculo unitario $x^2 + y^2 = 1$?

En cada uno de los ejercicios 15–19, se da una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ y un punto $\mathbf{p} \in U$. Compruebe que la derivada direccional de f en $\mathbf{p}$, en la dirección de (la tangente a) la curva de nivel que pasa por $\mathbf{p}$ (es decir, la curva $f(x, y) = f(\mathbf{p})$) es igual a cero.

15. $f(x, y) = 3x - y$, $\mathbf{p} = (2, 5)$

16. $f(x, y) = 5x^2 + 6y^2$, $\mathbf{p} = (-1, 0)$

17. $f(x, y) = xy$, $\mathbf{p} = (2, 3)$
18. $f(x, y) = \operatorname{sen}(x + y)$, $\mathbf{p} = (0, 0)$
19. $f(x, y) = e^x e^y$, $\mathbf{p} = (0, 0)$
20. Considere la función $f\colon \mathbb{R}^2 \to \mathbb{R}$ cuya gráfica es la superficie de revolución $z = f\left(\sqrt{x^2 + y^2}\right)$. Sea $\mathbf{p} = (x_0, y_0)$. Argumente por qué la derivada de f en $\mathbf{p}$ en la dirección del vector $\mathbf{u} = (-y_0, x_0)$ debe ser igual a cero. Ejemplifique este resultado con la función $f(x, y) = e^{-(x^2+y^2)}$. (Nota: los cálculos explícitos para este problema se podrán hacer cuando se haya estudiado cómo obtener las derivadas parciales de la función f, lo cual se hará en el siguiente capítulo).

Después de comprobar que las funciones dadas en los ejercicios 21–27 son homogéneas, verifique el teorema de Euler para este tipo de funciones, estudiado en el apéndice de esta sección (verifique tanto la fórmula del teorema como la de su corolario).

21. $f(x, y) = c$
22. $f(x, y) = ax + by$
23. $f(x, y) = x^2 + 5xy + y^2$
24. $f(x, y) = \operatorname{sen} \dfrac{x}{y}$
25. $f(x, y) = \dfrac{x + 3y}{2x - y}$
26. $f(x, y, z) = x^3 + y^3 + z^3 + xyz$
27. $f(x, y, z) = xy \ln \dfrac{xy + z^2}{x(y + z)}$
28. Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable, homogénea de grado α. Demuestre que las derivadas parciales de esta función $\frac{\partial f}{\partial x_i}\colon U \subseteq \mathbb{R}^n \to \mathbb{R}, i = 1, 2, \ldots, n$, son funciones homogéneas de grado $\alpha - 1$. Verifique este resultado con las funciones de los ejercicios 21–27.

2.8 Gradiente

Establezcamos formalmente la siguiente definición.

Definición. Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^n$. Se define el (vector) *gradiente* de la función f en el punto $\mathbf{x}_0 \in U$, denotado por grad $f(\mathbf{x}_0)$ o $\nabla f(\mathbf{x}_0)$, como el vector de $\mathbb{R}^n$ dado por

$$\operatorname{grad} f(\mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \frac{\partial f}{\partial x_2}(\mathbf{x}_0), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0)\right) \quad ■$$

Con este concepto, la fórmula (D) de la sección anterior, que nos da la derivada direccional de una función diferenciable $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ en el punto $\mathbf{x}_0 \in U$ en la dirección del vector unitario $\mathbf{v} \in \mathbb{R}^n$, se ve como

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \operatorname{grad} f(\mathbf{x}_0) \cdot \mathbf{v} \tag{A}$$

(la derivada de f en $\mathbf{x}_0$ en la dirección de $\mathbf{v}$ es el producto punto del vector gradiente de f en $\mathbf{x}_0$ con el vector $\mathbf{v}$).

Ejemplo 1. La función $f: \mathbb{R}^3 \to \mathbb{R}$ dada por $f(x, y, z) = x^2y^3z^4$ tiene por derivadas parciales a

$$\frac{\partial f}{\partial x} = 2xy^3z^4, \qquad \frac{\partial f}{\partial y} = 3x^2y^2z^4, \qquad \frac{\partial f}{\partial z} = 4x^2y^3z^3$$

En el punto (1, 1, 1) estas derivadas son

$$\frac{\partial f}{\partial x}(1, 1, 1) = 2, \qquad \frac{\partial f}{\partial y}(1, 1, 1) = 3, \qquad \frac{\partial f}{\partial z}(1, 1, 1) = 4$$

Entonces, el vector gradiente de la función f en el punto (1, 1, 1) es

$$\text{grad } f(1, 1, 1) = (2, 3, 4)$$

la derivada direccional de la función f en la dirección del vector unitario $\mathbf{v} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ es

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}}(1, 1, 1) &= \text{grad } f(1, 1, 1) \cdot \mathbf{v} = (2, 3, 4) \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \\ &= \frac{2}{\sqrt{3}} + \frac{3}{\sqrt{3}} + \frac{4}{\sqrt{3}} = \frac{9}{\sqrt{3}} = 3\sqrt{3}\end{aligned}$$ ■

Una de las cualidades importantes que tiene la fórmula (A) —además de su simplicidad—, es que permite descubrir algunos secretos importantes de las derivadas direccionales. En efecto, recordemos que el producto punto del vector $\mathbf{x} \in \mathbb{R}^n$ con el vector $\mathbf{y} \in \mathbb{R}^n$ tiene que ver con la proyección del vector $\mathbf{x}$ sobre el vector $\mathbf{y}$. En forma más precisa se tiene que la proyección de $\mathbf{x}$ sobre $\mathbf{y}$, denotada por $\text{PR}_{\mathbf{x}\to\mathbf{y}}$ es el vector dado por

$$\text{PR}_{\mathbf{x}\to\mathbf{y}} = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}$$

Más aún, si $\mathbf{y}$ es un vector unitario se tiene $\text{PR}_{\mathbf{x}\to\mathbf{y}} = (\mathbf{x} \cdot \mathbf{y})\mathbf{y}$. En el caso de la fórmula (A), poniendo $\mathbf{x} = \text{grad } f(\mathbf{x}_0)$, $\mathbf{y} = \mathbf{v}$ (que es un vector unitario), podemos concluir que

$$\text{PR}_{\text{grad } f(\mathbf{x}_0)\to\mathbf{v}} = (\text{grad } f(\mathbf{x}_0) \cdot \mathbf{v})\mathbf{v} = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)\mathbf{v}$$

de modo que entonces tenemos que la derivada direccional de la función diferenciable f en el punto $\mathbf{x}_0$ en la dirección del vector $\mathbf{v}$ es la componente del vector grad $f(\mathbf{x}_0)$ sobre el vector $\mathbf{v}$.

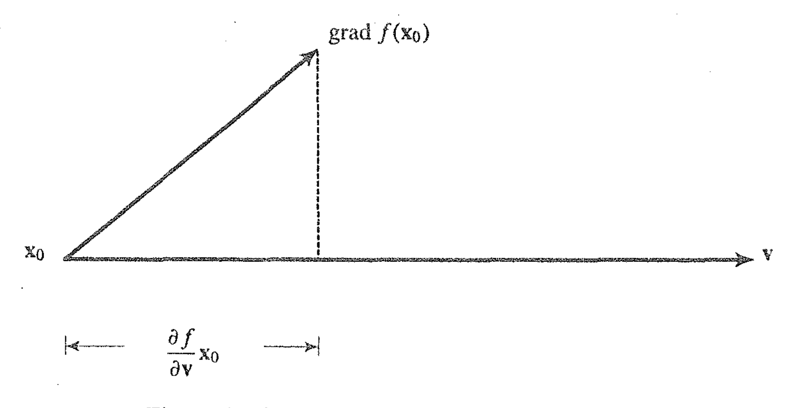


Figura 1. Proyección del gradiente sobre el vector v.

Este hecho nos permite obtener la siguiente importante conclusión: el valor de la derivada direccional $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ es máximo cuando el vector $\mathbf{v}$ está en la dirección del gradiente de f en $\mathbf{x}_0$ y nula cuando $\mathbf{v}$ es ortogonal a tal vector gradiente. Es decir, *el vector* grad $f(\mathbf{x}_0)$ *nos dice en qué dirección se tiene la mayor variación* (el mayor crecimiento) *de la función f en el punto* $\mathbf{x}_0$.

Esquemáticamente

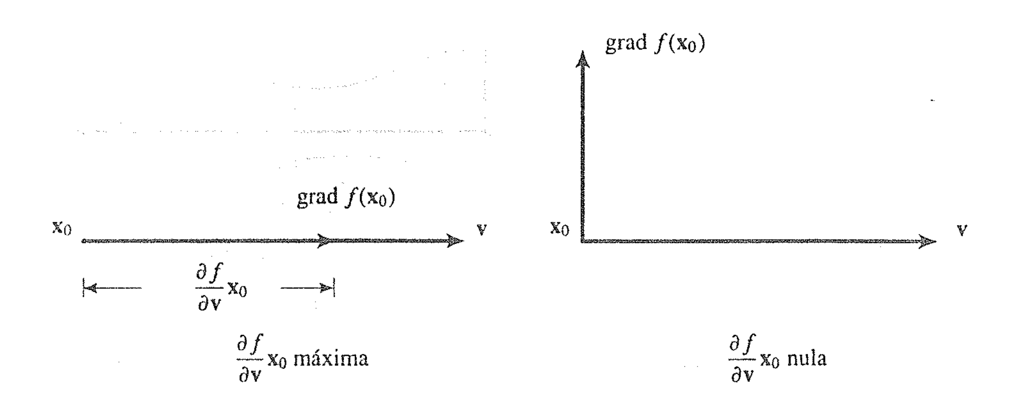


Figura 2. Los valores máximo y nulo de la derivada direccional.

Más aún, el valor máximo de esta derivada direccional es

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \|\operatorname{grad} f(\mathbf{x}_0)\| \qquad \text{donde } \mathbf{v} = \operatorname{grad} f(x_0)$$

lo cual se puede deducir, ya sea atendiendo a la figura anterior, o bien del hecho de que el ángulo θ entre el vector grad $f(x_0)$ y el vector unitario $\mathbf{v}$, es tal que

$$\operatorname{grad} f(\mathbf{x}_0) \cdot \mathbf{v} = \|\operatorname{grad} f(\mathbf{x}_0)\| \, \|\mathbf{v}\| \cos\theta = \|\operatorname{grad} f(\mathbf{x}_0)\| \cos\theta$$

de modo que cuando $\theta = 0$ (cuando grad $f(\mathbf{x}_0)$ y $\mathbf{v}$ son vectores en la misma dirección) se tiene

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \operatorname{grad} f(\mathbf{x}_0) \cdot \mathbf{v} = \|\operatorname{grad} f(\mathbf{x}_0)\| \cos 0 = \|\operatorname{grad} f(\mathbf{x}_0)\|$$

y cuando $\theta = \pi/2$ (cuando grad $f(\mathbf{x}_0)$ y $\mathbf{v}$ son vectores perpendiculares) se tiene

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \operatorname{grad} f(\mathbf{x}_0) \cdot \mathbf{v} = \|\operatorname{grad} f(\mathbf{x}_0)\| \cos\pi/2 = \operatorname{grad} f(\mathbf{x}_0)(0) = 0$$

También, retomando la serie de observaciones hechas en el ejemplo 3 de la sección anterior (que luego serán debidamente demostradas), en cuanto al hecho de que la derivada direccional de una función $z = f(x, y)$ en un punto $\mathbf{p}$ de alguna de sus curvas de nivel en la dirección de esta curva (de su tangente) es igual a cero, podemos ver que si $\mathbf{v}$ es un vector tangente a una curva de nivel de una función $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ en el punto $\mathbf{p} \in U$, entonces grad $f(\mathbf{p})$ debe ser un vector *ortogonal* a $\mathbf{v}$, pues en este caso se ha visto que la derivada direccional es cero. Así pues, tenemos que el *vector* grad $f(\mathbf{x}_0)$ es *un vector ortogonal* a la *curva de nivel que pasa por* $\mathbf{x}_0$. Esquemáticamente

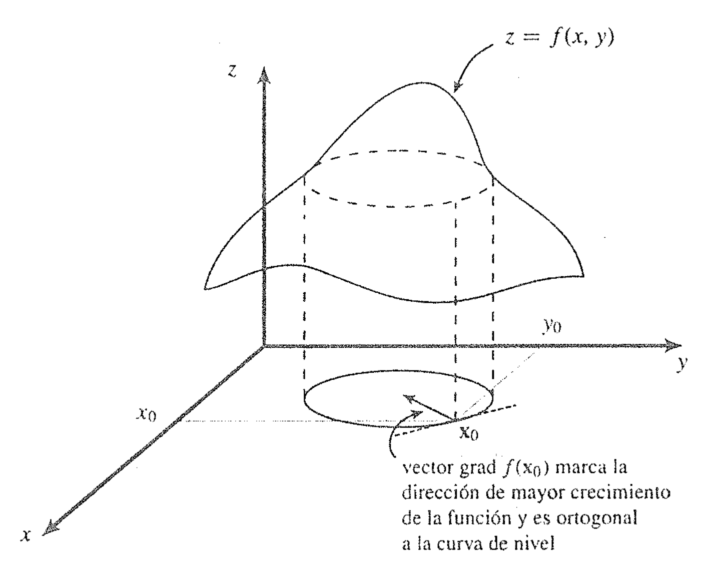


Figura 3. Los vectores grad $f(\mathbf{x})$ son ortogonales a las curvas de nivel.

Ejemplo 2. Considere la función $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = \sqrt{3 - x^2 - y^2}$. La superficie que representa la gráfica de esta función es la parte superior de la esfera de centro en el origen y radio $\sqrt{3}$. Si hacemos $f(x, y) = 1$, obtenemos la curva de nivel (el nivel 1) $x^2 + y^2 = 2$, que es una circunferencia con centro en el origen y radio $\sqrt{2}$. Las derivadas de la función f son

$$\frac{\partial f}{\partial x} = \frac{-x}{\sqrt{3 - x^2 - y^2}} \qquad \frac{\partial f}{\partial y} = \frac{-y}{\sqrt{3 - x^2 - y^2}}$$

En el punto $x_0 = (1, 1)$ —el cual es un punto de la curva de nivel $x^2 + y^2 = 2$— estas derivadas son

$$\frac{\partial f}{\partial x}(\mathbf{x}_0) = \frac{-1}{\sqrt{3 - 1 - 1}} = -1 \qquad \frac{\partial f}{\partial y}(\mathbf{x}_0) = \frac{-1}{\sqrt{3 - 1 - 1}} = -1$$

de modo que grad $f(\mathbf{x}_0) = (-1, -1)$. Este vector marca la dirección de mayor crecimiento de la función f en $\mathbf{x}_0$. La magnitud de esta derivada direccional máxima es $\|\text{grad } f(x_0)\| = \|(-1, -1)\| = \sqrt{2}$. Además, teniendo en cuenta que $\mathbf{u} = (-1, 1)$ es un vector tangente a la curva de nivel en $\mathbf{x}_0$, vemos que $\mathbf{u} \cdot \text{grad } f(1, 1) = (-1, 1) \cdot (-1, -1) = 1 - 1 = 0$; es decir, el vector grad $f(1, 1)$ es ortogonal a la curva de nivel en $\mathbf{x}_0$ (figura 4).

Ejemplo 3. Suponga que la distribución de la temperatura dentro de una habitación está dada por $T(x, y, z) = 25 + 0.02e^{0.1x+0.4y+0.01z^2}$ donde x, y, z se miden a partir de uno de los rincones (dado). A partir de ese rincón (el punto (0, 0, 0)), se quiere saber en qué dirección aumenta la temperatura con más rapidez. Según ya hemos visto, la velocidad de variación máxima se encuentra en la dirección del vector gradiente de T en (0, 0, 0), es decir, en la dirección del vector

$$\left(\frac{\partial T}{\partial x}(0, 0, 0), \frac{\partial T}{\partial y}(0, 0, 0), \frac{\partial T}{\partial z}(0, 0, 0)\right)$$

Calculando estas derivadas tenemos

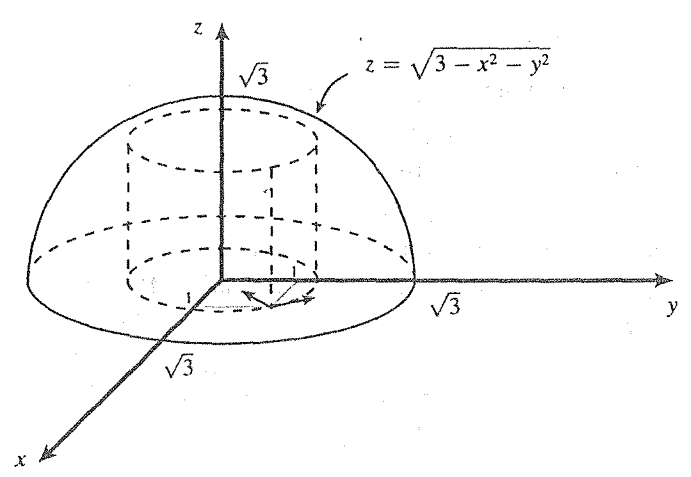


Figura 4. Ejemplo 2.

$$\frac{\partial T}{\partial x} = 0.002e^{0.1x+0.4y+0.01z^2}, \qquad \frac{\partial T}{\partial y} = 0.008e^{0.1x+0.4y+0.01z^2}$$

$$\frac{\partial T}{\partial z} = 0.0004ze^{0.1x+0.4y+0.01z^2}$$

Evaluando estas derivadas en $(0, 0, 0)$ obtenemos el vector grad $T(0, 0, 0) = (0.002, 0.008, 0)$ que nos marca la dirección de mayor crecimiento de la temperatura partiendo del punto de origen. ■

Ejemplo 4. Un cultivo de bacterias ha sido infectado por un contaminante, cuya concentración, medida con un sistema coordenado xy, está dada por

$$C = 2 + 4\,\text{sen}^2(x + 3y + 8xy)$$

donde x y y se miden en centímetros y C en centigramos por litro. Una bacteria b se encuentra en el punto de coordenadas $(0, 2)$. El vector grad $C(0, 2) = \left(\frac{\partial C}{\partial x}(0, 2), \frac{\partial C}{\partial y}(0, 2)\right) = (68\,\text{sen}\,12, 12\,\text{sen}\,12) = 4\,\text{sen}\,12(17, 3)$ nos da la dirección de movimiento de la bacteria en la cual ésta encontraría la mayor variación de concentración de contaminante. Para que la bacteria b se pueda mover en el cultivo sin tener cambio en la concentración del contaminante, lo tendrá que hacer sobre una curva de nivel de la función $C(x, y)$ correspondiente al nivel $C(0, 2) = 2 + 4\,\text{sen}^2 6$. ■

Las propiedades del vector gradiente ejemplificadas aquí con funciones de 2 variables, valen en general para funciones de n variables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. En particular quisiéramos hacer hincapié en la propiedad que establece que el vector grad $f(\mathbf{x}_0) \in \mathbb{R}^n$ es un vector "perpendicular" o "normal" a la curva (o más en general a la superficie) de nivel de la función f que pasa por $\mathbf{x}_0$.

Ejercicios (Capítulo 2, Sección 8)

1. Determine el vector gradiente de las siguientes funciones en los puntos indicados.

a. $f(x, y) = 3x^2y + \cos(xy)$ en el punto $\mathbf{p} = (1, 1)$

b. $f(x, y) = x^y$ en el punto $\mathbf{p} = (2, 2)$

c. $f(x, y) = \arctan \dfrac{x+y}{x-y}$ en el punto $(1, 0)$

d. $f(x, y, z) = (\operatorname{sen} x)(\cos y)(\tan z)$ en el punto $\mathbf{p} = (\pi/4, \pi/4, \pi/4)$

2. Calcule el ángulo entre los gradientes de la función

a. $f(x, y) = 3x^2 + y^2$ en los puntos $(2, 1)$ y $(1, 2)$,

b. $f(x, y, z) = xy^2z^3$ en los puntos $(1, 1, 1)$ y $(1, -1, -1)$.

3. Calcule el ángulo entre los gradientes de las funciones

a. $f(x, y) = xy$ y $g(x, y) = x^2 - y^2$ en el punto $(2, 1)$,

b. $f(x, y, z) = x^4 + 3y^4z$ y $g(x, y, z) = x + 3y - 2z$ en el punto $(1, 2, 1)$.

En los ejercicios 4–10 se da una función, un punto $\mathbf{p}$ de su dominio y un vector $\mathbf{u}$. Escriba en cada caso la descomposición del vector grad $f(\mathbf{p})$ en sus componentes ortogonales, una de las cuales está en la dirección del vector $\mathbf{u}$ dado. Es decir, determine el escalar α tal que grad $f(\mathbf{p}) = \alpha\mathbf{u} + \mathbf{w}$, donde $\mathbf{w}$ es un vector ortogonal a $\mathbf{u}$.

4. $f(x, y) = 4x - 16y$, $\mathbf{p} = (2, 1)$, $\mathbf{u} = (1, 1)$

5. $f(x, y) = x^2 + 3xy + y^2$, $\mathbf{p} = (1, 1)$, $\mathbf{u} = (0, 1)$

6. $f(x, y) = x \operatorname{sen} y + y \operatorname{sen} x$, $\mathbf{p} = (0, \pi)$, $\mathbf{u} = (2, 3)$

7. $f(x, y) = \dfrac{1+x^2}{1+y^2}$, $\mathbf{p} = (0, 0)$, $\mathbf{u} = (1, 1)$

8. $f(x, y, z) = x + y + z$, $\mathbf{p} = (1, 1, 1)$, $\mathbf{u} = (1, 1, 1)$

9. $f(x, y, z) = x^2 + y^2 + z^2$, $\mathbf{p} = (1, 1, 1)$, $\mathbf{u} = (-1, 0, 1)$

10. $f(x, y, z) = \ln(2 + x^2 + y^4 + z^6)$, $\mathbf{p} = (1, 2, 0)$, $\mathbf{u} = (0, 0, 5)$

11. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$, y sea $\mathbf{p} \in U$. Suponga que $\frac{\partial f}{\partial x}(\mathbf{p}) = 0$, $\frac{\partial f}{\partial y}(\mathbf{p}) = 5$.

a. Calcule la derivada de f en $\mathbf{p}$ en la dirección del vector $\mathbf{u} = (3, 4)$.

b. Calcule la derivada de f en $\mathbf{p}$ en la dirección del vector $\mathbf{u} = (\pi^2, 0)$.

c. Sea C una curva de nivel de f que pasa por $\mathbf{p}$. Demuestre que si $ax + by + c = 0$ es la ecuación de la recta tangente a C en $\mathbf{p}$, entonces $a = 0$.

d. Demuestre que la función f "crece" (en $\mathbf{p}$) en la dirección de cualquier vector en $\mathbb{R}^2$ en los cuadrantes primero y segundo, y "decrece" (en $\mathbf{p}$) en la dirección de cualquier vector en $\mathbb{R}^2$ en los cuadrantes tercero y cuarto.

12. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$ y sea $\mathbf{p} \in U$. Suponga que $\frac{\partial f}{\partial x}(\mathbf{p}) = 3$, $\frac{\partial f}{\partial y}(\mathbf{p}) = 4$. ¿En qué dirección se tiene que $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = 2$?, ¿en qué dirección se tiene $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = 0$?, ¿en qué dirección se tiene $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = -5$? ¿Hay alguna dirección en la que $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = 6$?

13. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el abierto U de $\mathbb{R}^2$ y sea $\mathbf{p}$ un punto de U en el que las derivadas parciales de f son iguales a cero. Pruebe que, de hecho, las derivadas de f en $\mathbf{p}$ en *cualquier dirección* son iguales a cero. Es decir, del hecho de que las derivadas de f en $\mathbf{p}$ en las direcciones de $\mathbf{i} = (1, 0)$ y $\mathbf{j} = (0, 1)$ son cero, se puede deducir que las derivadas de f en $\mathbf{p}$ en cualquier otra dirección son también nulas. Pruebe que esto es falso si no asuminos la diferenciabilidad de f (al menos en $\mathbf{p}$).

14. Sea $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^3$ y sea $\mathbf{p} \in U$. Suponga que $\frac{\partial f}{\partial x}(\mathbf{p}) = 6$, $\frac{\partial f}{\partial y}(\mathbf{p}) = 0$, $\frac{\partial f}{\partial z}(\mathbf{p}) = 8$. Demuestre que $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = 10$ es el máximo valor que puede tomar la derivada direccional de f en $\mathbf{p}$, y que este se logra en la dirección del vector unitario $\mathbf{u} = (3/5, 0, 4/5)$. ¿Cuál es el mínimo valor (absoluto) que puede tomar $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p})$?, ¿en qué dirección se tiene este valor?

15. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$ y sea $\mathbf{p} \in U$. Suponga que $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{p}) = 3$, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = 2$, donde $\mathbf{u} = (1/\sqrt{2}, -1/\sqrt{2})$, $\mathbf{v} = (\sqrt{3}/2, 1/2)$. Calcule las derivadas parciales de f en $\mathbf{p}$.

16. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$ y sea $\mathbf{p} \in U$. Suponga que $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{p}) = a$, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p}) = b$, donde los vectores unitarios $\mathbf{u} = (x_1, y_1)$ y $\mathbf{v} = (x_2, y_2)$ son linealmente independientes. Demuestre que las derivadas parciales de f en $\mathbf{p}$ son

$$\frac{\partial f}{\partial x}(\mathbf{p}) = \frac{ay_2 - by_1}{x_1 y_2 - x_2 y_1}, \qquad \frac{\partial f}{\partial y}(\mathbf{p}) = \frac{bx_1 - ax_2}{x_1 y_2 - x_2 y_1}$$

17. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = x^2 + y^2$.

- **a.** Determine los puntos $(x, y) \in \mathbb{R}^2$ en que el gradiente de esta función forma una ángulo de $\pi/4$ con el vector $\mathbf{u} = (1, 1)$.
- **b.** Determine los puntos $(x, y) \in \mathbb{R}^2$ donde el gradiente de esta función tiene la misma dirección del vector $\mathbf{u} = (-3, -4)$.
- **c.** Determine los puntos $(x, y) \in \mathbb{R}^2$ en que el gradiente de esta función es perpendicular al vector $\mathbf{u} = (-1, 1)$.

18. Sea $f: \mathbb{R}^3 \to \mathbb{R}$ la función $f(x, y, z) = z - x^2 - y^2$.

- **a.** Determine los puntos $(x, y, z) \in \mathbb{R}^3$ en que el gradiente de esta función forma un ángulo de $\pi/3$ con el vector $\mathbf{u} = (2, 1, 1)$.
- **b.** Determine los puntos $(x, y, z) \in \mathbb{R}^3$ donde el gradiente de esta función esté en la dirección del vector $\mathbf{u} = (1, 1, 1)$.
- **c.** Determine los puntos $(x, y, z) \in \mathbb{R}^3$ en que el gradiente de esta función es perpendicular al vector $\mathbf{u} = (2, -1, 1)$.

19. Demuestre que no existe función diferenciable alguna $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el abierto U de $\mathbb{R}^n$ para la que su derivada direccional en un punto $\mathbf{p} \in U$, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p})$, tenga signo constante en la dirección de *todo* vector unitario $\mathbf{v} \in \mathbb{R}^n$. ¿Puede ocurrir que $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p})$ sea igual a cero para todo vector unitario $\mathbf{v} \in \mathbb{R}^n$?

20. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable definida en el abierto U de $\mathbb{R}^n$ y sea $\mathbf{p} \in U$. Si $\mathbf{v}$ es un vector unitario en $\mathbb{R}^n$, demuestre que

$$\frac{\partial f}{\partial(-\mathbf{v})}(\mathbf{p}) = -\frac{\partial f}{\partial \mathbf{v}}(\mathbf{p})$$

21. Sean $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones diferenciables definidas en el conjunto abierto U de $\mathbb{R}^n$ y sea $\mathbf{p} \in U$. Demuestre las siguiente fórmulas para el gradiente

- **a.** $\operatorname{grad}(f + g)(\mathbf{p}) = \operatorname{grad} f(\mathbf{p}) + \operatorname{grad} g(\mathbf{p})$
- **b.** $\operatorname{grad}(fg)(\mathbf{p}) = f(\mathbf{p}) \operatorname{grad} g(\mathbf{p}) + g(\mathbf{p}) \operatorname{grad} f(\mathbf{p})$
- **c.** $\operatorname{grad}\left(\frac{f}{g}\right)(\mathbf{p}) = \dfrac{g(\mathbf{p}) \operatorname{grad} f(\mathbf{p}) - f(\mathbf{p}) \operatorname{grad} g(\mathbf{p})}{g^2(\mathbf{p})}$, si $g(\mathbf{p}) \neq 0$.

22. Pruebe que si $f: \mathbb{R}^n \to \mathbb{R}$ es una función constante, entonces grad $f(\mathbf{x}) = \mathbf{0}$ (el vector cero de $\mathbb{R}^n$) $\forall \mathbf{x} \in \mathbb{R}^n$. Más aún, suponga que $f: \mathbb{R}^n \to \mathbb{R}$ es una función definida en todo $\mathbb{R}^n$ tal que sus derivadas parciales son nulas en todo punto $\mathbf{x} \in \mathbb{R}^n$. Demuestre que f es constante.

23. Generalizando el resultado del ejercicio anterior: (demuestre que) la función $f: \mathbb{R}^n \to \mathbb{R}$ es constante si y sólo si las derivadas direccionales $\frac{\partial f}{\partial \mathbf{v}_i}(\mathbf{p})$ (existen y) son iguales a cero en todo punto $\mathbf{p} \in \mathbb{R}^n$, en las direcciones de un conjunto de n vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ linealmente independientes.

24. Considere las funciones $f(x, y) = 3x^2 + 2y^2$, $g(x, y) = 7 \ln x + \sqrt{3}y$. Demuestre que la derivada de la función f en el punto $\mathbf{p} = (1, 1)$ en la dirección del gradiente de la función g en $\mathbf{p}$ es igual a la derivada de la función g en $\mathbf{p}$ en la dirección del gradiente de la función f en $\mathbf{p}$. ¿Ocurre lo mismo con las funciones $f(x, y) = x^2 + y^2$, $g(x, y) = 2x + y$, en el punto $\mathbf{p} = (2, 1)$?

25. Sean $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones diferenciables definidas en el conjunto abierto U de $\mathbb{R}^n$. Sea $\mathbf{p}$ un punto de U. Demuestre que una condición necesaria y suficiente para que la derivada de f en $\mathbf{p}$ en la dirección del gradiente de g en $\mathbf{p}$ sea igual a la derivada de g en $\mathbf{p}$ en la dirección del gradiente de f en $\mathbf{p}$, es que las normas de los gradientes de f en $\mathbf{p}$ y de g en $\mathbf{p}$ sean iguales. Es decir, demuestre que

$$\frac{\partial f}{\partial \operatorname{grad} g(\mathbf{p})}(\mathbf{p}) = \frac{\partial g}{\partial \operatorname{grad} f(\mathbf{p})}(\mathbf{p}) \Leftrightarrow \|\operatorname{grad} f(\mathbf{p})\| = \|\operatorname{grad} g(\mathbf{p})\|$$

26. (Un ejercicio sobre elipses). Una elipse se define en geometría analítica clásica como el lugar geométrico de los puntos del plano tales que la suma de distancias a dos puntos fijos, llamados focos, es una magnitud constante. Sean $F_1 = (-c, 0)$ y $F_2 = (c, 0)$ los focos, y considere las funciones $f_1, f_2: \mathbb{R}^2 \to \mathbb{R}$

$$f_1(x, y) = \sqrt{(x + c)^2 + y^2} = \text{ distancia del punto } (x, y) \text{ a } F_1$$
$$f_2(x, y) = \sqrt{(x - c)^2 + y^2} = \text{ distancia del punto } (x, y) \text{ a } F_2$$

Sea $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = f_1(x, y) + f_2(x, y)$.

a. Describa las curvas de nivel de la función f.

b. Sea $\mathbf{u} \in \mathbb{R}^2$ un vector unitario tangente a una elipse en un punto de ella, digamos $\mathbf{p} = (x, y)$, apuntando en la dirección contraria a la del recorrido de las manecillas del reloj. Dé argumentos geométricos que muestren que $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{p}) = 0$. Deduzca de este hecho que

$$\operatorname{grad} f_1(x, y) \cdot \mathbf{u} + \operatorname{grad} f_2(x, y) \cdot \mathbf{u} = 0$$

c. Considere los vectores $\mathbf{v}_1 = -(x + c, y)$, $\mathbf{v}_2 = -(x - c, y)$ que unen el punto $\mathbf{p} = (x, y)$ de una elipse con los focos F_1 y F_2, respectivamente. Sea θ_1 el ángulo que forma el vector $\mathbf{v}_1$ con el vector $\mathbf{u}$ del inciso anterior, y sea θ_2 el ángulo que forma el vector $\mathbf{v}_2$ con el vector $-\mathbf{u}$. Demuestre que

$$\cos\theta_1 = \frac{1}{f_1(x, y)} \mathbf{v}_1 \cdot \mathbf{u} = \operatorname{grad} f_1(x, y) \cdot \mathbf{u}$$
$$\cos\theta_2 = -\frac{1}{f_2(x, y)} \mathbf{v}_2 \cdot \mathbf{u} = -\operatorname{grad} f_2(x, y) \cdot \mathbf{u}$$

d. Use los resultados de los dos incisos anteriores para concluír que $\theta_1 = \theta_2$. Obtenga entonces la conocida propiedad de la elipse: los radios que unen a un punto p de la elipse con los focos F_1 y F_2 forman ángulos iguales con la recta tangente a la elipse en el punto $\mathbf{p}$.

2.9 Vectores normales

Consideremos la gráfica de una función diferenciable $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$. Como sabemos, ésta es una superficie S en el espacio $\mathbb{R}^3$. Sea $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$ un punto de esta superficie. Queremos obtener una expresión que nos de un vector *normal* a la superficie S en $\mathbf{p}$. En este momento entenderemos la palabra "normal" ó "perpendicular" que califica la propiedad del vector procurado, tratando de extender lo que se entiende con este calificativo para gráficas de funciones de una sola variable (donde "normal" significa "perpendicular a la recta tangente"), sólo que ahora, en el caso de nuestra superficie S, debemos pensar en un "plano tangente a la superficie en $\mathbf{p}$" (lo que buenamente —de modo intuitivo— podemos pensar acerca de este término, que será objeto de estudio de la próxima sección). De modo pues que el vector que procuramos —el vector normal a la superficie S en $\mathbf{p}$—, será un vector perpendicular al plano tangente a la superficie en $\mathbf{p}$. (Veremos luego que, siendo la función f diferenciable, la existencia de este plano tangente está garantizada).

Llamemos $\mathbf{N_p} = (a, b, c)$ al vector normal procurado (ciertamente a, b y c dependen de $\mathbf{p}$). Siendo este vector perpendicular al plano tangente a S que pasa por $\mathbf{p}$, debe ser, en particular, perpendicular a cualquier recta tangente a la superficie S en $\mathbf{p}$. Ahora debemos recordar lo estudiado en la sección 2.5 sobre derivadas parciales: éstas nos daban las pendientes de las rectas tangentes —en las direcciones canónicas— a la superficie en $\mathbf{p}$.

Obtengamos dos vectores (en el plano xz y yz) que marquen las direcciones de las rectas tangentes a la superficie S en $\mathbf{p}$ (que sean paralelos a tales rectas).

Fijemos nuestra atención en el plano $y = y_0$. Este corta a la superficie S y da lugar a una curva —en ese plano— que en el punto $\mathbf{p}$ tiene una recta tangente de pendiente $\frac{\partial f}{\partial x}(x_0, y_0)$. Entonces el vector

$$\mathbf{v}_1 = \left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right)$$

es un vector en la dirección de la recta tangente mencionada (ver ejercicio 39 de la sección 3).

Esquemáticamente (Fig. 1)

NOTA: El vector $\mathbf{v}_1$ mostrado en la figura es en realidad el vector $\mathbf{v}_1 = \left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right)$ recorrido rígidamente al plano $y = y_0$.

De forma análoga, el vector

$$\mathbf{v}_2 = \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0)\right)$$

es un vector en la dirección de la recta tangente a la curva $z = f(x_0, y)$ en el punto $\mathbf{p}$ —la cual se encuentra en el plano $x = x_0$— (ver ejercicio 40 de la sección 3).

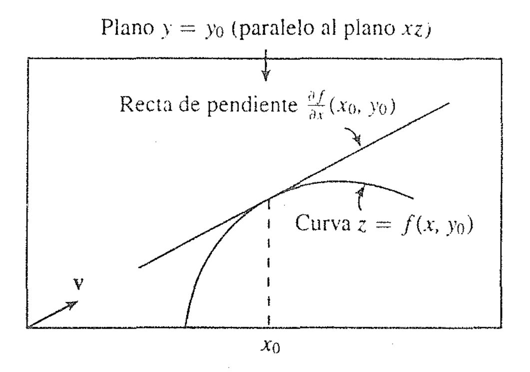


Figura 1. Gráfica de la recta tangente a la curva de intersección.

Así pues, el vector normal $\mathbf{N_p}$ procurado debe ser perpendicular a $\mathbf{v}_1$ y $\mathbf{v}_2$. Lo obtenemos como el producto cruz de $\mathbf{v}_1$ y $\mathbf{v}_2$, digamos $\mathbf{N_p} = \mathbf{v}_1 \times \mathbf{v}_2$. Tenemos

$$\begin{aligned}\mathbf{N_p} = \mathbf{v}_1 \times \mathbf{v}_2 &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial f}{\partial x}(x_0, y_0) \\ 0 & 1 & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} \\ &= \left(-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1\right) \end{aligned} \tag{1}$$

Esta es la expresión que nos permite obtener un vector perpendicular a la superficie dada por la gráfica de la función $z = f(x, y)$ en el punto $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$.

Ejemplo 1. Considere la función $f(x, y) = \sqrt{1 - x^2 - y^2}$. Como sabemos, su gráfica es la parte superior de la esfera con centro en el origen y radio 1. Obtengamos un vector perpendicular en el punto (x, y, z) cualquiera de ella. Todo lo que tenemos que hacer es calcular las derivadas parciales de la función, las cuales son

$$\frac{\partial f}{\partial x} = -\frac{x}{\sqrt{1 - x^2 - y^2}}, \qquad \frac{\partial f}{\partial y} = -\frac{y}{\sqrt{1 - x^2 - y^2}}$$

de modo que el vector normal es

$$\mathbf{N_p} = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right) = \left(\frac{x}{\sqrt{1 - x^2 - y^2}}, \frac{y}{\sqrt{1 - x^2 - y^2}}, 1\right)$$

o bien

$$\mathbf{N_p} = \frac{1}{\sqrt{1 - x^2 - y^2}}\left(x, y, \sqrt{1 - x^2 - y^2}\right)$$

Más aún, sabiendo que $z = \sqrt{1 - x^2 - y^2}$, vemos que un vector normal a tal esfera en el punto (x, y, z) es $\mathbf{N_p} = (x, y, z)$. ■

NOTA: Se ha venido hablando de *un* vector normal a una superficie en un punto. Es claro que, teniendo un vector $\mathbf{N_p}$ de esta naturaleza, se tiene automáticamente una infinidad de ellos, a saber

λN_p, con $\lambda \in \mathbb{R}$. Así pues, en el ejemplo anterior el vector normal que se obtuvo fue $\frac{1}{z}(x, y, z)$, el cual multiplicado por z es $N_p = (x, y, z)$. Es interesante notar también cómo en este ejemplo se juega la dualidad de papeles de los elementos del espacio $\mathbb{R}^3$ que ya se había comentado en el capítulo 1: por una parte, (x, y, z) juega el papel de un punto de la esfera y por otra parte el mismo (x, y, z) juega el papel de un vector normal a la esfera en ese punto (que es él mismo). Por razones "psicológicas" uno "distingue" estas dos ideas en una figura poniendo

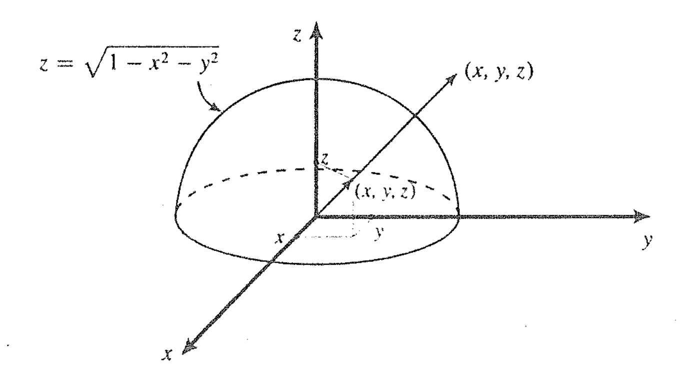


Figura 2. Gráfica de un punto y un vector normal a la esfera $\sqrt{1 - x^2 - y^2}$.

Ejemplo 2. Para la función $f(x, y) = x^2 + y^2$ que representa un paraboloide que abre hacia arriba, un vector normal en un punto (x, y, z) es

$$N_p = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right) = (-2x, -2y, 1)$$

Así por ejemplo, en (0, 0, 0) un vector normal es el (0, 0, 1). ■

Ejemplo 3. Para la función $f(x, y) = ax + by + c$ que representa un plano en el espacio, un vector normal al plano es

$$N_p = (-a, -b, 1)$$ ■

Quisiéramos ahora llamar la atención hacia un punto de especial interés que conecta lo discutido en esta sección con la propiedad del vector gradiente de ser perpendicular a una curva de nivel.

Comencemos por tener presente que una curva en el plano que sea la gráfica de la función $y = \varphi(x)$, también se puede ver como *una curva de nivel, correspondiente al nivel cero*, de la función $f(x, y) = y - \varphi(x)$. De hecho, los puntos del plano que constituyen la gráfica de la función $y = \varphi(x)$ son precisamente aquellos puntos que satisfacen que $y - \varphi(x) = 0$, es decir, en los cuales la función $f(x, y)$ es constante (igual a cero).

Así por ejemplo, la recta $y = x$, gráfica de la función $\varphi(x) = x$, es el nivel cero de la función $f(x, y) = y - x$, que geométricamente es un plano que pasa por el origen (figura 3).

Nótese que al decir "nivel cero", nos estamos refiriendo, geométricamente, al *corte* de la superficie $z = f(x, y)$ con el plano xy. Más aún, si tuviéramos una curva en el plano, no necesariamente la gráfica de una función $y = \varphi(x)$, también podremos en *ciertas condiciones* (la pregunta sería ¿cuándo podremos?) verla como el nivel cero de alguna función $z = f(x, y)$, es decir, como la manera como

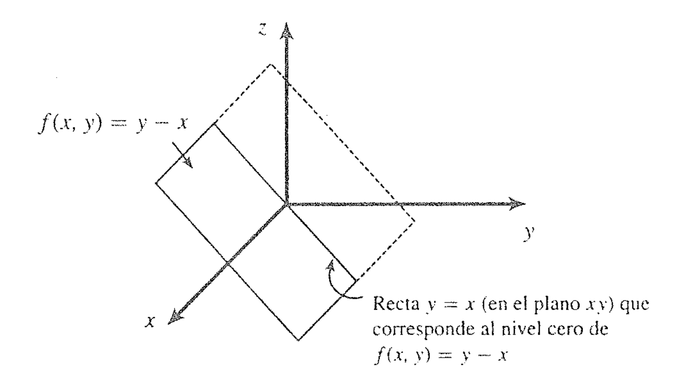


Figura 3. Gráfica del plano $f(x, y) = y - x$.

esta superficie corta al plano xy. Esta es una cuestión que será discutida ampliamente en el próximo capítulo (seción 4). Por ejemplo, la curva $x^2 + y^2 = 1$ no es (globalmente) la gráfica de función $y = \varphi(x)$ alguna. Sin embargo, la podemos ver como el nivel cero de la función $f(x, y) = x^2 + y^2 - 1$ (esta función no es única: para cualquier constante $k \neq 0$, la función $f(x, y) = k(x^2 + y^2 - 1)$ tiene por nivel cero a $x^2 + y^2 = 1$; esto significa que existen muchas superficies en $\mathbb{R}^3$ que cortan el plano xy en la curva $x^2 + y^2 = 1$), cuya gráfica es un paraboloide que abre hacia arriba y cuyo vértice está en el punto $(0, 0, -1)$ (figura 4).

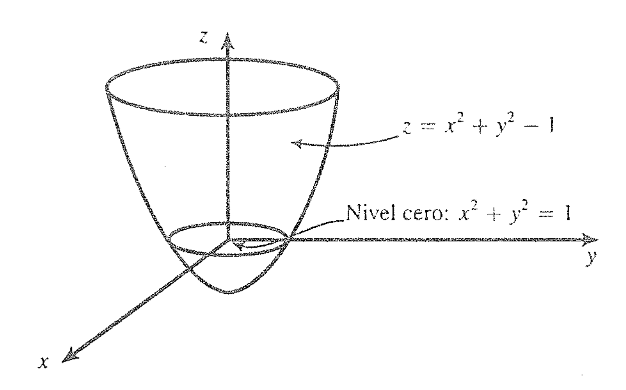


Figura 4. Gráfica del paraboloide $z = x^2 + y^2 - 1$ y su nivel cero.

Llevando esta idea una dimensión más arriba, tenemos que la superficie que representa la gráfica de la función $z = f(x, y)$ también la podemos ver como el nivel cero de la función de tres variables $F(x, y, z) = z - f(x, y)$. Entonces, la superficie $z = f(x, y)$ se puede ver como una *superficie de nivel* (correspondiente al nivel cero) de la función $F(x, y, z) = z - f(x, y)$, o bien, como la manera en que la gráfica de la función $F(x, y, z) = z - f(x, y)$ (que vive en el espacio $\mathbb{R}^4$) "atraviesa" el espacio $\mathbb{R}^3$ (identificado como una porción de $\mathbb{R}^4$, correspondiente a aquellos puntos de $\mathbb{R}^4$ cuya última coordenada sea igual a cero). Así por ejemplo, la superficie $z = x^2 + y^2$ es la superficie de nivel de

la función $F(x, y, z) = z - x^2 - y^2$, correspondiente al nivel 0 de esta función. Más aún, podríamos tener una superficie en el espacio $\mathbb{R}^3$, que no representa (globalmente) la gráfica de alguna función $z = f(x, y)$ y también verla como una superficie de nivel de una función $u = F(x, y, z)$. Este es el caso que se presenta con la esfera $x^2 + y^2 + z^2 = 1$, la cual no es la gráfica de función $z = f(x, y)$ alguna, pero sí la podemos ver como el nivel cero de la función $F(x, y, z) = x^2 + y^2 + z^2 - 1$.

Así las cosas, si vemos la superficie $z = f(x, y)$ como el nivel cero de la función $F(x, y, z) = z - f(x, y)$, y recordamos que el gradiente de esta función es un vector perpendicular a cualquier superficie de nivel de ella, concluimos que el vector

$$\operatorname{grad} F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)$$

debe ser normal a la superficie considerada. En efecto, si calculamos estas derivadas parciales obtenemos

$$\operatorname{grad} F = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right)$$

resultado que coincide con la fórmula (1) establecida anteriormente.

Ejemplo 4. Retomemos el ejemplo 1. La superficie $z = f(x, y) = \sqrt{1 - x^2 - y^2}$ es un trozo de la esfera $x^2 + y^2 + z^2 = 1$. Lo que digamos entonces de la esfera, vale, en particular, para la gráfica de $z = f(x, y)$. Esta esfera la podemos ver como la superficie de nivel (cero) de la función $F(x, y, z) = x^2 + y^2 + z^2 - 1$ cuyo gradiente es

$$\operatorname{grad} F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right) = (2x, 2y, 2z) = 2(x, y, z)$$

Este vector debe ser entonces normal a la superficie. En particular $\mathbf{N_p} = (x, y, z)$ lo es, como ya observamos en el ejemplo 1. ■

Ejemplo 5. Consideremos el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Para obtener un vector normal en el punto $\mathbf{p} = (x, y, z)$ de él, lo podemos ver como superficie de nivel de la función $F: \mathbb{R}^3 \to \mathbb{R}$, dada por

$$F(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$$

El vector grad F debe ser entonces normal a esta superficie de nivel. Este vector es

$$\operatorname{grad} F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right) = \left(\frac{2x}{a^2}, \frac{2y}{b^2}, \frac{2z}{c^2}\right) = 2\left(\frac{x}{a^2}, \frac{y}{b^2}, \frac{z}{c^2}\right)$$

En particular entonces, el vector $\mathbf{N_p} = \left(\frac{x}{a^2}, \frac{y}{b^2}, \frac{z}{c^2}\right)$ es un vector normal al elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ en un punto (x, y, z) de él. ■

Ejercicios (Capítulo 2, Sección 9)

Para cada una de las funciones $z = f(x, y)$ dadas en los ejercicios 1–10, determine un vector normal a su gráfica en el punto indicado.

1. $f(x, y) = 1$ en un punto cualquiera $\mathbf{p} = (x_0, y_0)$
2. $f(x, y) = -128\pi^2$ en un punto cualquiera $\mathbf{p} = (x_0, y_0)$
3. $f(x, y) = x + 12$ en un punto cualquiera $\mathbf{p} = (x_0, y_0)$
4. $f(x, y) = x - 125$ en un punto cualquiera $\mathbf{p} = (x_0, y_0)$
5. $f(x, y) = xy$ en un punto cualquiera $\mathbf{p} = (x_0, y_0)$
6. $f(x, y) = \operatorname{sen} x + \operatorname{sen} y$ en el punto $\mathbf{p} = (0, 0)$
7. $f(x, y) = e^y \cos x$ en el punto $\mathbf{p} = (0, 1)$
8. $f(x, y) = 2x^2 + 5y^3$ en el punto $\mathbf{p} = (2, -1)$
9. $f(x, y) = \ln(2 + x + y)$ en el punto $\mathbf{p} = (0, 0)$
10. $f(x, y) = \operatorname{sen}(\operatorname{sen} x \cos y)$ en el punto $\mathbf{p} = (\pi, \pi)$

En cada uno de los ejercicios 11–15, se dan ecuaciones de superficies en $\mathbb{R}^3$. Con ellas como superficies de nivel de funciones $u = F(x, y, z)$, determine un vector normal a cada una en los puntos indicados.

11. $xyz = 1$ en el punto $\mathbf{p} = (1, 1, 1)$
12. $xy + xz + yz - 3 = 0$ en el punto $\mathbf{p} = (1, 1, 1)$
13. $x^2 + y^2 - z^2 = 0$ en los puntos $\mathbf{p} = (3, 4, 5)$ y $\mathbf{q} = (-3, -4, -5)$
14. $x^2y^2 + x^2z^2 + y^2z^2 + xyz - 4 = 0$ en el punto $\mathbf{p} = (1, 1, 1)$
15. $x^y + x^z + z^x - 3xyz = 0$ en el punto $\mathbf{p} = (1, 1, 1)$
16. ¿Cierto o falso? La gráfica de una función $z = f(x, y)$ no puede tener por vector normal a un vector en el plano xy. Explique.
17. ¿Cierto o falso? Si la gráfica de la función $z = f(x, y)$ tiene al vector $N \in \mathbb{R}^3$ por vector normal en el punto $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$, entonces N no puede ser normal a la gráfica de la función en otro punto $\mathbf{q}$ de ella, diferente de $\mathbf{p}$. ¿Cambia la respuesta si consideramos una superficie más general en $\mathbb{R}^3$ (no necesariamente la gráfica de una función $z = f(x, y)$), por ejemplo, una superficie cerrada (como una esfera o un elipsoide)?
18. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$, y sea $\mathbf{p} = (x_0, y_0) \in U$. Llamemos $\mathbf{N_p}$ al vector normal a la gráfica de f en $(x_0, y_0, f(x_0, y_0))$. El vector grad $f(\mathbf{p})$ es un vector en $\mathbb{R}^2$, que se puede identificar con un vector en $\mathbb{R}^3$ con su última coordenada nula. Demuestre que el ángulo θ entre los vectores $\mathbf{N_p}$ y grad $f(\mathbf{p})$ está dado por

$$\theta = \arctan \frac{1}{\|\operatorname{grad} f(x_0, y_0)\|}$$

19. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = e^{-(x^2+y^2)}$. Sea θ el ángulo del ejercicio anterior. Demuestre que $\lim_{\|(x,y)\| \to \infty} \theta = \pi/2$. Interprete geométricamente este hecho.

En los ejercicios 20–25 se da una función $z = f(x, y)$ y un vector $\mathbf{N} \in \mathbb{R}^3$. Determine el (los) punto(s) de la gráfica de la función (si los hay) para los que el vector $\mathbf{N}$ es un vector normal.

20. $f(x, y) = x^2 + y^2$, $\mathbf{N} = (0, 0, 1)$

21. $f(x, y) = x + y - 3$, $\mathbf{N} = (1, 1, -1)$

22. $f(x, y) = 3x - 2y + 4$, $\mathbf{N} = (4, 5, 1)$

23. $f(x, y) = 2x^2 + 3xy + 5y^2$, $\mathbf{N} = (3, 2, -3)$

24. $f(x, y) = \ln(1 + x + 2y)$, $\mathbf{N} = (-1, -3, 4)$

25. $f(x, y) = \operatorname{sen}\sqrt{x^2 + y^2}$, $\mathbf{N} = (0, 0, -3)$

En los ejercicios 26–30 se da la ecuación de una superficie S y un vector $\mathbf{N} \in \mathbb{R}^3$. Determine el (o los) punto(s) de S (si los hay) para los que $\mathbf{N}$ es un vector normal a S.

26. $x^2 + y^2 + z^2 = 4$, $\mathbf{N} = (2, 2, 2)$

27. $x^2 + 2y^2 + 3z^2 = 1$, $\mathbf{N} = (-2, 3, 6)$

28. $x^2 + 4y^2 - z^2 = 1$, $\mathbf{N} = (0, 3, 4)$

29. $2x^2 + 2y^2 - 5z^2 + 2xy - 2x - 4y - 4z - 11 = 0$, $\mathbf{N} = (5, 4, 3)$

30. $7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 18 = 0$, $\mathbf{N} = (1, 0, 1)$

31. Considere la función diferenciable $f: \mathbb{R} \to \mathbb{R}$, $z = f(y)$. Suponga que ésta tiene un máximo local en el punto $y = y_0 > 0$, $z = z_0 > 0$. Sea $z = f\left(\sqrt{x^2 + y^2}\right)$ la superficie de revolución generada al poner a girar la gráfica de $z = f(y)$ alrededor del eje z. ¿Para qué puntos de la superficie el vector $\mathbf{N} = (0, 0, 1)$ es un vector normal? Ejemplifique este resultado con la función $f(y) = y^2 e^{-y^2}$. (Nota: la pregunta puede ser contestada en base a consideraciones geométricas. Para convencerse analíticamente de su respuesta, puede hacer uso de las siguientes fórmulas de las derivadas parciales de la función $z = f\left(\sqrt{x^2 + y^2}\right)$, las cuales se entenderán después de estudiar el siguiente capítulo

$$\frac{\partial z}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} f'(u), \qquad \frac{\partial z}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}} f'(u)$$

donde $u = \sqrt{x^2 + y^2}$).

2.10 Planos tangentes

Consideremos la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$. Supongamos que es diferenciable en el punto $(x_0, y_0) \in U$. Queremos asociar a ella (a su gráfica) un plano "tangente" que haga el papel de la recta tangente a la gráfica de una función de una sola variable. Por supuesto que esto nos llevaría a contestar la pregunta: ¿qué se entiende por un plano *tangente* a una superficie en un punto? La respuesta a esta pregunta está intimamente ligada con la propiedad de diferenciabilidad de la función en el punto, como veremos más adelante. Por lo pronto, pensemos que el plano que buscamos debe poder contener a todas las rectas tangentes a la superficie $z = f(x, y)$

en el punto $(x_0, y_0, f(x_0, y_0))$. Más aún, el plano que se procura es tal que si cortamos a la superficie $z = f(x, y)$ con un plano perpendicular al plano xy que pase por $(x_0, y_0, 0)$, entonces la recta tangente a la curva de intersección de la superficie con este plano, en el punto $(x_0, y_0, f(x_0, y_0))$, debe estar contenida en el "plano tangente" que buscamos.

Entonces, en particular, tal plano debe tener por vector normal al vector $\mathbf{N_p} = \left(-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1\right)$, según se discutió en la sección anterior. Por lo tanto, sabiendo que $\mathbf{N_p}$ es un vector normal al plano buscado y que éste debe pasar (digamos que debe "tocar") por el punto $(x_0, y_0, f(x_0, y_0))$, concluimos que su ecuación debe ser

$$-\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + z - f(x_0, y_0) = 0$$

o sea

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \tag{T}$$

Observando esta ecuación con detenimiento nos encontramos con que todo lo que se necesita para construir la ecuación del plano tangente a la gráfica de la función $z = f(x, y)$ en el punto $\mathbf{p} = (x_0, y_0)$ son las derivadas parciales de esta función calculadas en $\mathbf{p}$. ¿Quiere decir esto que si una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ posee derivadas parciales en el punto (x_0, y_0), ya podemos asociarle un plano tangente a su gráfica en el punto $(x_0, y_0, f(x_0, y_0))$?

Para dar respuesta a esta pregunta nos vemos obligados (de nuevo) a establecer con precisión el concepto de tangencia de un plano con una superficie. De manera "formal" la respuesta a la pregunta sería "sí": hallamos las derivadas parciales y las sustituimos en la ecuación anteriormente obtenida y ... ¡ya está! Ese es nuestro plano tangente procurado. Sin embargo, la experiencia vivida en la sección 4 (ver ejemplo 6) nos dice que la función $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

de modo que $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, tendría por plano tangente en el origen a $z = 0$ (plano xy), aun siendo la gráfica de la función discontinua en $(0, 0)$ (situación sumamente desagradable, pues la idea intuitiva de "tangencia" que tenemos exige *al menos* la continuidad de la gráfica en el punto donde el plano toca a la superficie).

Acontece que el concepto de tangencia que funciona en matemáticas es "el mismo" que el concepto de diferenciabilidad de funciones. De modo más preciso, si la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en (x_0, y_0) entonces diremos que la ecuación (T) *define* al plano *tangente* a la superficie $z = f(x, y)$ en el punto $(x_0, y_0, f(x_0, y_0))$. Es decir, es cuando la función f es diferenciable (y sólo en este caso) cuando el plano dado por la ecuación (T) es *tangente* (en el sentido preciso que se entiende en matemáticas) a la gráfica de la función en el punto considerado.

Veamos cuál es el sentido geométrico de la "tangencia" del plano dado por (T). Siendo la función f diferenciable en (x_0, y_0) tenemos que

$$f(x_0 + h_1, y_0 + h_2) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h_1 + \frac{\partial f}{\partial y}(x_0, y_0)h_2 + r(h_1, h_2)$$

donde

$$\lim_{(h_1,h_2)\to(0,0)} \frac{r(h_1, h_2)}{\|(h_1, h_2)\|} = 0$$

Poniendo $h_1 = x - x_0$, $h_2 = y - y_0$, nos queda la expresión anterior convertida en

$$f(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + r(x - x_0, y - y_0)$$

donde

$$\lim_{(x,y)\to(x_0,y_0)} \frac{r(x - x_0, y - y_0)}{\|(x - x_0, y - y_0)\|} = 0$$

Al ver esta expresión nos encontramos con la agradable sorpresa de que la suma de los 3 primeros términos del segundo miembro nos da precisamente (según la ecuación (T)) la altura, en el punto (x, y), del plano tangente a $z = f(x, y)$ en (x_0, y_0). De modo que el residuo $r(x - x_0, y - y_0)$ se puede ver como

$$r(x - x_0, y - y_0) = f(x, y) - \left(f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right)$$

= diferencia entre la z de la función $z = f(x, y)$ en el punto (x, y) y la z en el mismo punto del plano tangente a la gráfica en (x_0, y_0)

El hecho de que f sea una función diferenciable nos dice que este residuo es *muy pequeño* en torno al punto (x_0, y_0). Con más precisión, nos dice que si lo dividimos entre la norma del vector $(x - x_0, y - y_0)$, este cociente va a cero cuando x tiende a x_0 y y tiende a y_0 [Es claro que el puro residuo $r(x - x_0, y - y_0)$ tiende a cero cuando $(x, y) \to (x_0, y_0)$. Sin embargo esta propiedad sólo nos dice que la función f es continua en (x_0, y_0), como ya lo habíamos advertido anteriormente —¿por qué?]. Esta propiedad del residuo es precisamente la que le da el *carácter de tangencia* al plano dado por (T). Es decir, es justamente cuando el residuo $r(x - x_0, y - y_0)$ (visto como la diferencia de la z de $z = f(x, y)$ y la z del plano dado por (T)) tiene tal propiedad (que al dividirlo por la norma de $(x - x_0, y - y_0)$ tiende a cero cuando $(x, y) \to (x_0, y_0)$), cuando decimos que el plano dado por (T) es *tangente* a la superficie $z = f(x, y)$ en (x_0, y_0).

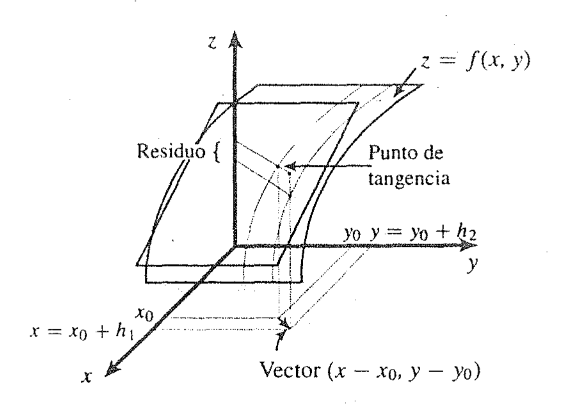


Figura 1. Plano tangente a la gráfica de $z = f(x, y)$.

Como las funciones con las que trabajamos normalmente son diferenciables, entonces podremos hablar con libertad de planos tangentes a sus gráficas. Veamos ahora algunos ejemplos.

Ejemplo 1. Se quiere hallar la ecuación del plano tangente a la superficie $z = x^2 + y^2$ en los puntos $(0, 0, 0)$ y $(1, 2, 5)$. Todo lo que necesitamos son las derivadas parciales $\frac{\partial z}{\partial x} = 2x$ y $\frac{\partial z}{\partial y} = 2y$. En $x = y = 0$, tenemos $\frac{\partial z}{\partial x}(0, 0) = 0$, $\frac{\partial z}{\partial y}(0, 0) = 0$, de modo que según la ecuación (T), el plano tangente es $z = 0$, es decir, el plano xy (resultado que coincide con la respuesta intuitiva al imaginar el paraboloide $z = x^2 + y^2$ cuyo vértice es justamente el origen). En el punto en que $x = 1$, $y = 2$, tenemos $\frac{\partial z}{\partial x}(1, 2) = 2$, $\frac{\partial z}{\partial y}(1, 2) = 4$, de modo que en este punto el plano tangente es $z = 5 + 2(x - 1) + 4(y - 2)$, o sea $2x + 4y - z = 5$. ■

Consideremos una superficie S, digamos que dada como la gráfica de la función diferenciable $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$. Se define la *recta normal* a la superficie S en el punto $\mathbf{p} = (x_0, y_0, z_0)$ de ella, como la recta que pasa por $\mathbf{p}$ y contiene el (o es paralela al) vector normal a la superficie en $\mathbf{p}$ (es decir, la recta perpendicular al plano tangente a la superficie en $\mathbf{p}$). Según lo discutido en esta sección y en la anterior, tal recta debe de ser paralela al vector $\mathbf{N_p} = \left(-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1\right)$ y, como debe también pasar por el punto $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$, su ecuación debe ser

$$\begin{cases} x = x_0 - \dfrac{\partial f}{\partial x}(x_0, y_0)t \\ y = y_0 - \dfrac{\partial f}{\partial y}(x_0, y_0)t \\ z = f(x_0, y_0) + t \end{cases} \qquad t \in \mathbb{R}$$

Ejemplo 2. Se quiere hallar la ecuación del plano tangente y de la recta normal a la superficie $z = x^2y + e^{x^2+y^2}$ en el punto en que $x = 1$, $y = 1$. Tenemos que

$$\frac{\partial z}{\partial x} = 2xy + 2xe^{x^2+y^2}, \qquad \frac{\partial z}{\partial y} = x^2 + 2ye^{x^2+y^2}$$

de modo que con $x = 1$, $y = 1$ nos queda que $z = 1 + e^2$, y

$$\frac{\partial z}{\partial x}(1, 1) = 2 + 2e^2, \qquad \frac{\partial z}{\partial y}(1, 1) = 1 + 2e^2$$

La ecuación del plano tangente es entonces

$$z = 1 + e^2 + (2 + 2e^2)(x - 1) + (1 + 2e^2)(y - 1)$$

o bien

$$(2 + 2e^2)x + (1 + 2e^2)y - z = 2 + 3e^2$$

La ecuación de la recta normal es

$$\begin{cases} x = 1 - (2 + 2e^2)t \\ y = 1 - (2 + 2e^2)t \\ z = 1 + e^2 + t \end{cases} \qquad t \in \mathbb{R}$$

■

Ejemplo 3. Consideremos el paraboloide $z = x^2 + y^2$. Sea $\mathbf{p} = (x_0, y_0, z_0)$ un punto cualquiera de él. Como $\frac{\partial z}{\partial x} = 2x$ y $\frac{\partial z}{\partial y} = 2y$, tenemos que la ecuación de la recta normal a esta superficie en $\mathbf{p}$ es

$$\begin{cases} x = x_0 - 2x_0t \\ y = y_0 - 2y_0t \\ z = x_0^2 + y_0^2 + t \end{cases} \qquad t \in \mathbb{R}$$

Nótese que si ponemos $t = \frac{1}{2}$, se tiene $x = y = 0$. Esto significa que la recta normal corta al eje z. Como el punto $\mathbf{p}$ es arbitrario, tenemos probada la siguiente propiedad geométrica del paraboloide $z = x^2 + y^2$: todas sus rectas normales cortan el eje z. En realidad esta es una propiedad válida para superficies más generales (superficies de revolución), hecho que se considera en el ejercicio 42 al final de esta sección. ■

Ejemplo 4. Consideremos el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Como sabemos, esta función no es la gráfica de función alguna del tipo $z = f(x, y)$. Sin embargo, esto no nos impide obtener la ecuación del plano tangente a ella en un punto $\mathbf{p} = (x_0, y_0, z_0)$. Todo lo que necesitamos es un vector normal a la superficie en $\mathbf{p}$. En el ejemplo 5 de la sección anterior, vimos que un vector tal es $\mathbf{N_p} = \left(\frac{x_0}{a^2}, \frac{y_0}{b^2}, \frac{z_0}{c^2}\right)$, de modo que el plano tangente al elipsoide en $\mathbf{p}$ es

$$\frac{x_0}{a^2}(x - x_0) + \frac{y_0}{b^2}(y - y_0) + \frac{z_0}{c^2}(z - z_0) = 0$$

o bien

$$\frac{x_0}{a^2}x + \frac{y_0}{b^2}y + \frac{z_0}{c^2}z = \frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} + \frac{z_0^2}{c^2} = 1$$

(la última igualdad se justifica por ser $\mathbf{p}$ un punto del elipsoide). Es decir, la ecuación del plano tangente al elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ en un punto $\mathbf{p} = (x_0, y_0, z_0)$ de él es

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} + \frac{z_0 z}{c^2} = 1$$ ■

Ejemplo 5. Tomemos de nuevo el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Sea $\alpha x + \beta y + \gamma z = \delta$ un plano dado en el espacio. Nos preguntamos por el plano tangente al elipsoide que sea paralelo al plano dado. En realidad, hay dos respuestas a esta pregunta, pues localizando un plano tangente al elipsoide (digamos que "por arriba") podemos construir siempre otro tangente a él y paralelo al anterior (digamos que "por abajo"). Nótese que la propiedad de los planos $A_1x + B_1y + C_1z = D_1$ y $A_2x + By_2 + Cz_2 = D_2$ de ser paralelos significa que sus vectores normales son paralelos, es decir que hay una constante $k \neq 0$ tal que $(A_1, B_1, C_1) = k(A_2, B_2, C_2)$. En nuestro caso, debemos hallar *en qué puntos* del elipsoide sus planos tangentes tienen vectores normales paralelos al vector (α, β, γ), que es el vector normal al plano dado (y por tanto es un vector normal a los planos tangentes que procuramos). Como ya sabemos, un vector normal al elipsoide en el punto (x_0, y_0, z_0) viene dado por $\left(\frac{x_0}{a^2}, \frac{y_0}{b^2}, \frac{z_0}{c^2}\right)$, de modo que debemos resolver para x_0, y_0, z_0 la ecuación

$$\left(\frac{x_0}{a^2}, \frac{y_0}{b^2}, \frac{z_0}{c^2}\right) = k(\alpha, \beta, \gamma)$$

donde k es alguna constante. Es claro que los valores de x_0, y_0, z_0 que satisfacen tal condición son $x_0 = k\alpha a^2$, $y_0 = k\beta b^2$, $z_0 = k\gamma c^2$. Por otra parte, el punto (x_0, y_0, z_0) debe ser punto del elipsoide dado, por lo que

$$\begin{aligned} 1 &= \frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} + \frac{z_0^2}{c^2} = \frac{(k\alpha a^2)^2}{a^2} + \frac{(k\beta b^2)^2}{b^2} + \frac{(k\gamma c^2)^2}{c^2} \\ &= k^2(\alpha^2 a^2 + \beta^2 b^2 + \gamma^2 c^2) \end{aligned}$$

de donde

$$k = \pm\frac{1}{\sqrt{\alpha^2 a^2 + \beta^2 b^2 + \gamma^2 c^2}}$$

Llamando $\eta = \sqrt{\alpha^2 a^2 + \beta^2 b^2 + \gamma^2 c^2}$, tenemos que hay dos posibilidades para k, a saber $k = \pm\eta^{-1}$. Entonces, uno de los puntos del elipsoide que tiene plano tangente paralelo al dado es $(x_0, y_0, z_0) = \eta^{-1}(\alpha a^2, \beta b^2, \gamma c^2)$, en cuyo caso el plano tangente es

$$\alpha(x - \eta^{-1}\alpha a^2) + \beta(y - \eta^{-1}\beta b^2) + \gamma(z - \eta^{-1}\gamma c^2) = 0$$

o sea

$$\alpha x + \beta y + \gamma z = \eta^{-1}(\alpha^2 a^2 + \beta^2 b^2 + \gamma^2 c^2) = \eta^{-1}\eta^2 = \eta$$

Se obtiene fácilmente que el otro plano tangente paralelo al plano dado es $\alpha x + \beta y + \gamma z = -\eta$. Así pues, los planos tangentes al elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ que son paralelos al plano $\alpha x + \beta y + \gamma z = \delta$ son

$$\alpha x + \beta y + \gamma z = \pm\sqrt{\alpha^2 a^2 + \beta^2 b^2 + \gamma^2 c^2}$$ ■

Ejemplo 6. La gráfica de la función de una variable $f: \mathbb{R} - \{0\} \to \mathbb{R}$ dada por $y = \frac{a^2}{x}$ (una hipérbola con asíntotas los ejes coordenados) tiene una propiedad muy interesante: si trazamos la recta tangente en cualquier punto de ella, ésta formará con los ejes coordenados un triángulo cuya área será constante (es decir, no depende del punto escogido).

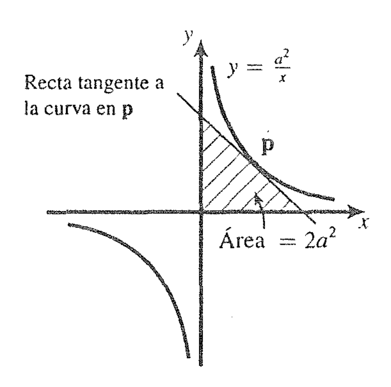


Figura 2. El triángulo formado por la recta tangente a $y = a^2/x$ y los ejes coordenados tiene área constante.

Este hecho es fácil de probar y lo dejamos como ejercicio para el lector. En este ejemplo queremos ver una propiedad análoga de una superficie en el espacio. Si tomamos la función $f(x, y) = \frac{a^2}{xy}$ (definida en todo $\mathbb{R}^2$ excepto en los puntos (x, y) tales que $xy = 0$, i.e. en los planos coordenados xz y yz), se tendrá que el plano tangente a la superficie que ella representa formará un tetraedro de volumen constante con los planos coordenados. Demostrar este hecho es el objetivo de este ejemplo. Tomemos entonces un punto cualquiera $\mathbf{p} = (x_0, y_0, z_0)$ de la superficie $z = \frac{a^2}{xy}$ y escribamos la ecuación del plano tangente que pasa por $\mathbf{p}$. Las derivadas parciales son

$$\frac{\partial z}{\partial x} = -\frac{a^2}{x^2 y}, \qquad \frac{\partial z}{\partial y} = -\frac{a^2}{xy^2}$$

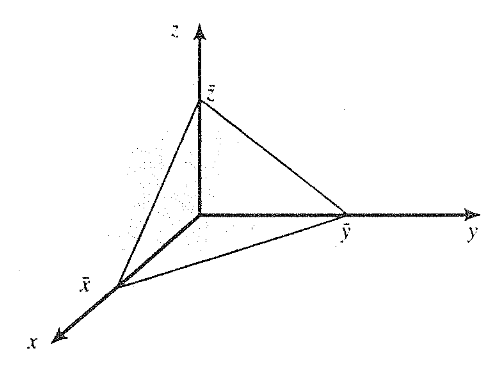


Figura 3. Tetraedro del ejemplo 6.

de modo que la ecuación del plano tangente procurado es

$$z = -\frac{a^2}{x_0^2 y_0}x - \frac{a^2}{x_0 y_0^2}y + 3\frac{a^2}{x_0 y_0}$$

Para hallar el volúmen del tetraedro que este plano forma con los planos coordenados, todo lo que necesitamos son las intersecciones que tiene con los ejes coordenados, digamos $\bar{x}$ (con el eje x), $\bar{y}$ (con el eje y) y $\bar{z}$ (con el eje z). En tal caso el volumen del tetraedro será $V = \frac{1}{6}\bar{x}\bar{y}\bar{z}$ (figura 3).

Haciendo $x = y = 0$, obtenemos $\bar{z}$ de la ecuación del plano tangente Nos queda que

$$\bar{z} = \frac{3a^2}{x_0 y_0}$$

Análogamente, haciendo $x = z = 0$, obtenemos $\bar{y} = 3y_0$, y haciendo $y = z = 0$ obtenemos $\bar{x} = 3x_0$.

Entonces el volumen del tetraedro es

$$V = \frac{1}{6}\bar{x}\bar{y}\bar{z} = \frac{1}{6}(3x_0)(3y_0)\left(\frac{3a^2}{x_0 y_0}\right) = \frac{9}{2}a^2$$

resultado que no depende de las coordenadas del punto **p**, como queríamos que ocurriera. ■

Ejemplo 7. Regresando de nuevo al caso de funciones de una variable, recordamos que la familia de curvas en el plano $x^2 + y^2 = ax$, $x^2 + y^2 = by$ (a y b son parámetros) se dicen "ortogonales", pues, en los puntos comunes, éstas se cortan ortogonalmente, es decir, sus rectas tangentes en tales puntos son perpendiculares entre sí. Nuevamente, invitamos al lector a que verifique esta propiedad. Geométricamente estas curvas se ven como (ambas son familias de circunferencias)

En el caso de superficies en el espacio, también podemos hablar de "superficies ortogonales", refiriéndonos con esto a que en los puntos donde las superficies coincidan, sus planos tangentes son ortogonales (o con más precisión, los vectores normales a tales superficies en los puntos comunes, son ortogonales entre sí). Por ejemplo, veamos que las familias de esferas (análogas a las circunferencias del plano) $x^2 + y^2 + z^2 = ax$, y, $x^2 + y^2 + z^2 = by$ son ortogonales. Nótese que los puntos comunes de las esferas de ambas familias son puntos $(x, y, z) \in \mathbb{R}^3$ que cumplen con la condición $ax = by$ (además, claro está, con las dos ecuaciones de las esferas). Sea **p** un tal punto. Obtengamos un

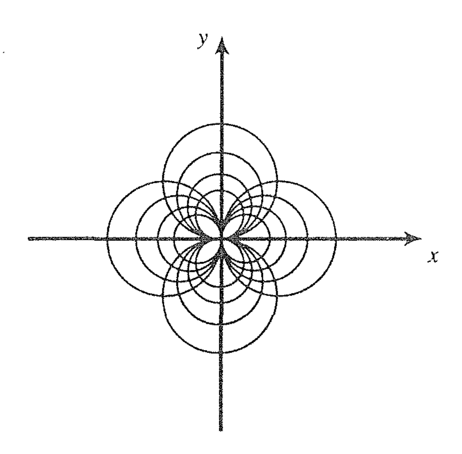


Figura 4. Circunferencias ortogonales $x^2 + y^2 = ax$, $x^2 + y^2 = by$.

vector normal a cada una de las esferas (que comparten a $\mathbf{p}$) en $\mathbf{p}$. Esto lo podemos hacer calculando los gradientes de las funciones

$$F_1(x, y, z) = x^2 + y^2 + z^2 - ax, \qquad F_2(x, y, z) = x^2 + y^2 + z^2 - by$$

cuyo nivel cero son precisamente las esferas de nuestro problema. Tenemos

$$\begin{aligned} \text{grad } F_1 &= (2x - a, 2y, 2z) \\ \text{grad } F_2 &= (2x, 2y - b, 2z) \end{aligned}$$

Queremos ver que en el punto $\mathbf{p}$ estos vectores son ortogonales. Esto lo logramos convenciéndonos que su producto punto es cero. En efecto

$$\begin{aligned} \text{grad } F_1 \cdot \text{grad } F_2 &= (2x - a, 2y, 2z) \cdot (2x, 2y - b, 2z) \\ &= 4x^2 - 2ax + 4y^2 - 2by + 4z^2 \\ &= 4(x^2 + y^2 + z^2) - 2ax - 2by \end{aligned}$$

en el punto $\mathbf{p}$ se tiene $by = ax$, y, $x^2 + y^2 + z^2 = ax$, de modo que

$$\text{grad } F_1 \cdot \text{grad } F_2 = 4ax - 2ax - 2(ax) = 0$$

como queríamos que ocurriera. ■

Ejercicios (Capítulo 2, Sección 10)

En los ejercicios 1–5 se da una función diferenciable $z = f(x, y)$ y un punto $\mathbf{p} = (x_0, y_0)$ de su dominio. Determine la ecuación del plano tangente a la gráfica de la función dada en el punto $\mathbf{p}$.

1. $f(x, y) = 3x + 8y - 10$, $\mathbf{p} = (x_0, y_0)$

2. $f(x, y) = x^3 + 8y^3$, $\mathbf{p} = (0, 0)$

3. $f(x, y) = x^y$, $\mathbf{p} = (2, 1)$

4. $f(x, y) = x \arctan(\arctan y)$, $\mathbf{p} = (1, 0)$

5. $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$, $\mathbf{p} = (0, 0)$

6. Demuestre que si la superficie S en $\mathbb{R}^3$ es una superficie de nivel de la función diferenciable $u = f(x, y, z)$, y si $\mathbf{p} = (x_0, y_0, z_0) \in S$ es un punto en el que las derivadas parciales $\frac{\partial F}{\partial x}(\mathbf{p})$, $\frac{\partial F}{\partial y}(\mathbf{p})$, $\frac{\partial F}{\partial z}(\mathbf{p})$ no son simultáneamente cero, entonces la ecuación del plano tangente a S en $\mathbf{p}$ viene dada por

$$\frac{\partial F}{\partial x}(\mathbf{p})(x - x_0) + \frac{\partial F}{\partial y}(\mathbf{p})(y - y_0) + \frac{\partial F}{\partial z}(\mathbf{p})(z - z_0) = 0$$

7. Considere la función $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^2 + y^2 - z^2$. Ciertamente ésta es una función diferenciable en todo punto $\mathbf{p} \in \mathbb{R}^3$. En particular en el origen lo es. Su nivel cero es una superficie S determinada. ¿Existe un plano tangente para S en el origen? Discuta ampliamente (por ejemplo, la relación entre "diferenciabilidad" y "existencia de planos tangentes", el papel que juegan las hipótesis hechas sobre las derivadas parciales de la función F en el ejercicio anterior, etc.).

8. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^2$, sea $\mathbf{p} \in U$ y sea $\mathbf{v} = (a, b) \in \mathbb{R}^2$ un vector unitario. En el ejercicio 24 de la sección 5 se probó que la curva de intersección de la superficie $z = f(x, y)$ con un plano perpendicular al plano xy que pasa por $\mathbf{p}$, paralelo al vector $\mathbf{v}$, tiene por recta tangente en $\mathbf{p}$ a

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = f(x_0, y_0) + \dfrac{\partial f}{\partial \mathbf{v}}(x_0, y_0)t \end{cases} \qquad t \in \mathbb{R}$$

Demuestre que, para todos los posibles vectores $\mathbf{v} \in \mathbb{R}^2$, las rectas tangentes anteriores se encuentran en el plano tangente a $z = f(x, y)$ en $\mathbf{p}$.

En los ejercicios 9–11 se da la ecuación de una superficie en el espacio tridimensional y un punto $\mathbf{p}$ de ella. Determine la ecuación del plano tangente a la superficie en el punto $\mathbf{p}$.

9. $z^2 + 3z - x^2 - y^2 - 2 = 0$, $\mathbf{p} = (1, 1, 1)$

10. $x - y^2 - z^2 = 0$, $\mathbf{p} = (0, 0, 0)$

11. $x^2 + y^2 + z^2 - 4x - 8y - 16z + 54 = 0$, $\mathbf{p} = (1, 2, 3)$

12. Determine la ecuación del plano tangente a la superficie $z = x^2 + y^2$ que sea paralelo al plano $3x + 8y - 5z = 10$.

13. El problema anterior puede ser resuelto "algebraicamente" (sin cálculo) de la siguiente manera:

   a. El plano tangente procurado debe ser de la forma $3x + 8y - 5z = k$, para alguna $k \in \mathbb{R}$. Explique.

b. Resuelva simultáneamente las ecuaciones $z = x^2 + y^2$, $3x + 8y - 5z = k$, para obtener que los puntos comunes del paraboloide y del plano deben satisfacer $5x^2 - 3x + 5y^2 - 8y + k = 0$.

c. Viendo a la expresión $5x^2 - 3x + 5y^2 - 8y + k = 0$ como una ecuación cuadrática para x, imponga la condición de que el plano tangente procurado debe compartir *un solo punto* con el paraboloide, para concluír que $(-3)^2 - 4(5)(5y^2 - 8y + k) = 0$, o sea $100y^2 - 160y + 20k - 9 = 0$.

d. Repita el argumento del inciso anterior con la expresión $100y^2 - 160y + 20k - 9 = 0$ para concluír que $(-160)^2 - 4(100)(20k - 9) = 0$, de donde se obtiene que $k = 73/20$. De aquí entonces que el plano tangente procurado es $3x + 8y - 5z = 73/20$.

14. Halle la ecuación de los planos tangentes a la superficie $3x^2 + 5y^2 + 3z^2 - 2xy + 2xz - 2yz - 12x = 0$ que sean paralelos al plano $z = 0$:

a. usando las técnicas del cálculo discutidas en esta sección;

b. usando la técnica del ejercicio anterior.

15. Verifique que no existen planos tangentes a la superficie $2x^2 + 2y^2 - 5z^2 + 2xy - 2x - 4y - 4z + 2 = 0$ que sean paralelos al plano $x = 0$.

16. Determine la ecuación del plano tangente a la superficie $z = x^2 + xy$ que sea perpendicular a los planos $x + y - z = 3$ y $2x - y + z = 4$. ¿Existe alguna técnica algebraica análoga a la del ejercicio 13 que permita resolver este problema?

17. Determine la ecuación del plano tangente a la superficie $z = 3x^2 - 8xy + 5y^2$ en el punto en que la recta normal tenga por vector paralelo a $\mathbf{v} = (-1, 0, 2)$.

18. Halle la ecuación del plano tangente a la superficie $z = x^2 + y^2 - 4x$ que sea perpendicular a la recta $x = 3 + 4t$, $y = -2t$, $z = 1 + t$, $t \in \mathbb{R}$.

19. Determine las ecuaciones de los planos tangentes al elipsoide $x^2 + 3y^2 + 5z^2 = 1$ que sean paralelos al plano tangente a la superficie $z = xy$ en el punto $\mathbf{p} = (1, 1, 1)$.

20. Considere la superficie $x^2 + y^2 + z^2 - 2x - 2y - 6z + 10 = 0$. Demuestre que esta superficie tiene una infinidad de planos tangentes perpendiculares a cualquier plano dado $ax + by + cz = d$. Determine la expresión general de estos planos. Explique geométricamente su resultado.

21. Dados dos planos perpendiculares $a_1x + b_1y + c_1z = d_1$, $a_2x + b_2y + c_2z = d_2$, demuestre que la superficie del ejercicio anterior tiene exactamente dos planos tangentes perpendiculares a los planos dados. Determine estos planos tangentes si los planos dados son $3x - y + z = 7$, $x + 8y + 5z = 10$. ¿Puede resolverse el problema si los planos dados no son perpendiculares? Explique.

22. Repita los dos ejercicios anteriores con la superficie $7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 30 = 0$.

En los ejercicios 23–26 determine los puntos de la superficie dada en los que los planos tangentes sean paralelos a los planos coordenados.

23. $x^2 + 5y^2 + 10z^2 = 12$

24. $(x - 2)^2 + 5(y - 3)^2 + 10(z + 1)^2 = 12$.

25. $x^2 + 3y^2 + z^2 + 4x - 6y - 2z + 7 = 0$.

26. $7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 18 = 0$

27. Hallar los puntos del elipsoide $x^2 + 2y^2 + 3z^2 = 6$ en los que la recta normal que pasa por ellos es perpendicular al plano $4x - 6y + 3z = 7$.

28. Determine las ecuaciones de los planos tangentes al elipsoide $x^2 + y^2 + 2z^2 = 2$ en los puntos de intersección de éste con la recta $x = 3t$, $y = 2t$, $z = t$, $t \in \mathbb{R}$.

29. Hallar las ecuaciones de los planos tangentes a la superficie $z = x^2 + 3y^2$ en los puntos de intersección de ésta con la recta que resulta de la intersección de los dos planos $2x - y - z = 0$, $x + 3y - 4z = 0$.

30. Hallar el volumen del tetraedro que forman los planos coordenados con el plano tangente al elipsoide $x^2 + y^2 + 3z^2 = 50$ en el punto $(1, 5, 2\sqrt{2})$

31. Determine los puntos del elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ en los que los planos tangentes cortan a los ejes coordenados en puntos equidistantes del origen.

32. Demuestre que los planos tangentes a la superficie $\sqrt{x} + \sqrt{y} + \sqrt{z} = \sqrt{a}$ cortan los ejes coordenados en puntos cuya suma de distancias al origen es constante.

33. Demuestre que las rectas normales a la superficie $z = \frac{1}{2}x^2 + \frac{1}{3}y^3 + y$, en los puntos (x_0, y_0, z_0), con $x_0 \neq 0$, no cortan al eje z.

Los ejercicios 34–41 tratan sobre planos tangentes y rectas normales a esferas. Cada uno de ellos puede ser resuelto con técnicas independientes del cálculo y, por supuesto, con técnicas del cálculo como las discutidas en esta sección. Invitamos al lector a que resuelva cada ejercicio de las dos maneras para que vea en cada caso las ventajas o desventajas que conlleva cada una de estas técnicas.

34. Demostrar que el plano $ax + by + cz + d = 0$ es tangente a la esfera $x^2 + y^2 + z^2 = r^2$ si y solamente si $a^2r^2 + b^2r^2 + c^2r^2 = d^2$.

35. Demostrar que el plano $2x - 6y + 3z - 49 = 0$ es tangente a la esfera $x^2 + y^2 + z^2 = 49$. ¿En qué punto? Hallar el otro plano tangente a la esfera que sea paralelo al dado.

36. Hallar las ecuaciones de los planos tangentes a la esfera $x^2 + y^2 + z^2 - 10x + 2y + 26z - 113 = 0$ que sean paralelos a las rectas

$$\begin{cases} x = 2t \\ y = -3t \\ z = 2t \end{cases} \qquad \begin{cases} x = 1 + 3t \\ y = -1 - 2t \\ z = t \end{cases}$$

37. Los puntos $A = (2, 5, 3)$ y $B = (-1, -2, -3)$ son los extremos de un diámetro de una esfera. Hallar las ecuaciones de los planos tangentes a esta esfera en los puntos A y B.

38. Una esfera tiene su centro en el punto $(3, 4, 5)$ y pasa por el origen de coordenadas. Hallar la ecuación del plano tangente a la esfera en el origen. Obtenga también el otro plano tangente a la esfera que sea paralelo al plano hallado.

39. Determine la esfera que tenga por planos tangentes a los planos paralelos $x + y + z = 5$, $x + y + z = -3$, sabiendo que el punto $(1, 2, 2)$ es uno de los puntos de contacto.

40. Hallar las ecuaciones de los planos tangentes a la esfera con centro en (x_0, y_0, z_0) y radio r, que sean paralelos al plano $ax + by + cz = d$.

41. Hallar las ecuaciones de los planos tangentes a la esfera $x^2 + y^2 + z^2 - 4x - 4y - 8z + 20 = 0$ que contengan:

a. al eje x,

b. al eje y,

c. al eje z.

42. Considere la gráfica de la función diferenciable $z = f(y)$, dibujada en el primer cuadrante del plano zy. Sea $z = f\left(\sqrt{x^2 + y^2}\right)$ la superficie de revolución que se obtiene al girar $z = f(y)$ alrededor del eje z. Ciertamente los puntos $\mathbf{p}_1 = (0, y_1, f(y_1))$, $\mathbf{p}_2 = (0, -y_1, f(y_1))$, $\mathbf{p}_3 = (y_1, 0, f(y_1))$ y $\mathbf{p}_4 = (-y_1, 0, f(y_1))$ son puntos de la superficie de revolución mencionada. Obtenga las ecuaciones de los planos tangentes y de las rectas normales a la superficie en cada uno de estos puntos. Compruebe que los cuatro planos se cortan en un mismo punto en el eje z, y que esto mismo ocurre con las rectas normales.

43. (Generalización del ejercicio anterior). Use las fórmulas de las derivadas parciales de la función $z = f\left(\sqrt{x^2 + y^2}\right)$ mencionadas en la nota del ejercicio 31 de la sección anterior, para demostrar que todas los planos tangentes en los puntos de la superficie $z = f\left(\sqrt{x^2 + y^2}\right)$ correspondientes a un mismo nivel (en los que $x^2 + y^2 =$ cte.), se cortan en un mismo punto sobre el eje z, y que todas las rectas normales a la superficie se cortan con el eje z (su eje de revolución).

44. Considere el cono $z = \sqrt{x^2 + y^2}$. Sea $\mathbf{p} = (x_0, y_0, \sqrt{x^2 + y^2})$ un punto distinto del origen. Demuestre que la ecuación del plano tangente al cono en el punto $\mathbf{p}$ es

$$z = \frac{1}{\sqrt{x^2 + y^2}}(x_0 x + y_0 y)$$

Demuestre también que este plano tangente "se intersecta" con el cono en los puntos de la recta

$$\begin{cases} x = x_0 t \\ y = y_0 t \\ z = \sqrt{x_0^2 + y_0^2}\, t \end{cases} \qquad t \in \mathbb{R}$$

la cual es una recta que pasa por el origen y tiene a $\mathbf{p}$ por vector paralelo. Verifique que, de hecho, todos los puntos de esta recta, excepto el origen, comparten el mismo plano tangente con $\mathbf{p}$.

45. En el ejemplo 4 se obtuvo que el plano tangente al elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ en el punto $\mathbf{p} = (x_0, y_0, z_0)$ es $\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} + \frac{z_0 z}{c^2} = 1$. con este resultado demuestre que los planos tangentes al elipsoide en sus puntos de intersección con el plano $z = z_0$ $(-c < z_0 < c,\ z_0 \neq 0)$ se cortan en un mismo punto sobre el eje z. Más aún, demuestre que éstos son los planos tangentes al cono elíptico

$$z = -\frac{c}{z_0}\sqrt{c^2 - z^2}\sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}} + \frac{c^2}{z_0}$$

(en todos sus puntos excepto en $\left(0, 0, \frac{c^2}{z_0}\right)$). En particular obtenga que la esfera $x^2 + y^2 + z^2 = a^2$ "está metida tangencialmente" dentro de la parte del cono $z = -\sqrt{3}\sqrt{x^2 + y^2} + 2a$ comprendida entre su vértice $(0, 0, 2a)$ y el plano $z = -a$.

46. Demuestre que las esferas $x^2 + y^2 + z^2 = 2ax$, $x^2 + y^2 + z^2 = 2by$, $x^2 + y^2 + z^2 = 2cz$ se cortan entre sí ortogonalmente.

2.11 La diferencial

Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable en el conjunto abierto U de $\mathbb{R}^2$. Entonces, para cada $\mathbf{x} \in U$ tenemos

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \frac{\partial f}{\partial x}h_1 + \frac{\partial f}{\partial y}h_2 + r(\mathbf{h})$$

donde

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

a la parte lineal en h_1 y h_2 de esta expresión se le llama *diferencial* de la función f en (x, y) y se denota por $df(x, y)$ o simplemente por df. Así

$$df = \frac{\partial f}{\partial x}h_1 + \frac{\partial f}{\partial y}h_2$$

Obsérvese que si $f(x, y) = x$, se obtiene $h_1 = dx$ y si $f(x, y) = y$ se obtiene $h_2 = dy$. Se escribe entonces

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

Retomando la definición de diferenciabilidad de la función f, podemos escribir

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + df(\mathbf{x}) + r(\mathbf{h})$$

o bien, como para $\mathbf{h}$ pequeño (es decir $\|\mathbf{h}\|$ pequeño), se tiene $r(\mathbf{h}) \approx \mathbf{0}$ (puesto que $\lim_{\mathbf{h}\to\mathbf{0}} r(\mathbf{h}) = \mathbf{0}$), se debe tener que

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) \approx df(\mathbf{x})$$

Geométricamente podemos interpretar la diferencial de la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ en $\mathbf{x} \in U$ como sigue (ver la figura 1 de la sección anterior): El vector $\mathbf{x}+\mathbf{h}$ está cerca del vector $\mathbf{x}$ (con $\mathbf{h}$ pequeño); la función f sufre un incremento (positivo, negativo o cero) al pasar su argumento de $\mathbf{x}$ a $\mathbf{x}+\mathbf{h}$. Tal incremento está dado por $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})$, el cual se puede ver formado por dos sumandos: uno de ellos es el residuo $r(\mathbf{h})$ que es la distancia entre el plano tangente a la superficie en $\mathbf{x}$ y la superficie misma, en el punto $\mathbf{x}+\mathbf{h}$, y el otro es justamente lo que acabamos de definir como la diferencial de la función en $\mathbf{x}$. Pensando en que $\mathbf{h}$ es pequeño (en norma), es que podemos decir que tal incremento se puede aproximar por la diferencial de la función.

Para el caso de funciones de una variable (donde se tienen representaciones geométricas más sencillas), estas ideas —que son completamente análogas— se ven geométricamente como (figura 1).

Ejemplo 1. Se quiere hacer un cálculo aproximado de $(1.08)^{(3.98)}$. Para esto, podemos usar las ideas anteriores y considerar la función $f(x, y) = x^y$ en el punto $(1, 4)$. Tomando $\mathbf{h}_1 = 0.08$ y $\mathbf{h}_2 = -0.02$, vemos que

$$f((1, 4) + (0.08, -0.02)) = (1.08)^{(3.98)}$$

en tanto que $f(1, 4) = 1^4 = 1$. Como

$$f((1, 4) + (0.08, -0.02)) - f(1, 4) \approx df(1, 4)$$

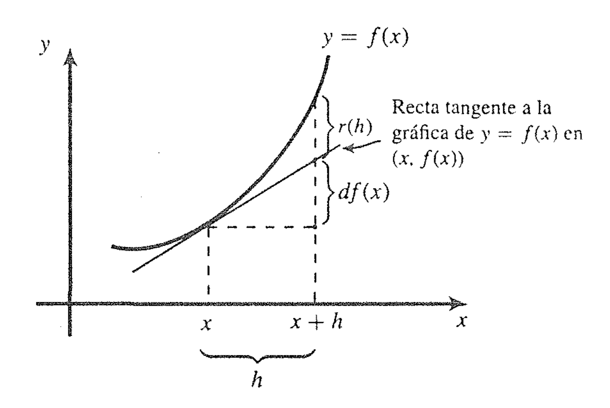


Figura 1. La diferencial df para funciones de una variable.

tenemos que si calculamos $df(1, 4)$ obtendremos una estimación de la cantidad $(1.08)^{(3.98)}$.

$$\begin{aligned} df(1,4) &= \frac{\partial f}{\partial x}(1,4)h_1 + \frac{\partial f}{\partial y}(1,4)h_2 \\ &= \left(yx^{y-1}\big|_{\substack{x=1\\y=4}}\right)h_1 + \left(x^y \ln x\big|_{\substack{x=1\\y=4}}\right)h_2 \\ &= 4(0.08) = 0.32 \end{aligned}$$

Así

$$(1.08)^{3.98} \approx 1 + 0.32 = 1.32$$

(el valor exacto es 1.358396). ■

Ejemplo 2. Supongamos que las funciones $f, g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son diferenciables. Entonces (Teorema 2.6.2) la función $fg: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ también es diferenciable. Obtengamos $d(fg)$. Se tiene

$$\begin{aligned} d(fg) &= \frac{\partial(fg)}{\partial x}dx + \frac{\partial(fg)}{\partial y}dy \\ &= \left(f\frac{\partial g}{\partial x} + g\frac{\partial f}{\partial x}\right)dx + \left(f\frac{\partial g}{\partial y} + g\frac{\partial f}{\partial y}\right)dy \\ &= f\left(\frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy\right) + g\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dx\right) \\ &= fdg + gdf \end{aligned}$$

De manera análoga podemos obtener sin dificultad las fórmulas

$$\begin{aligned} d(f+g) &= df + dg \\ d\left(\frac{f}{g}\right) &= \frac{gdf - fdg}{g^2} \text{ si } g(x) \neq 0, \ x \in U \end{aligned}$$

Se ve entonces que la diferencial de una función de dos variables tiene propiedades análogas a las de la derivada (y de la diferencial) de una función de una variable. ■

En el caso más general de una función de n variables $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable, tenemos que la diferencial de f se define como

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i$$

siendo también df una aproximación del incremento de la función $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})$ con $\mathbf{h}$ pequeño.

Ejemplo 3. Se quiere calcular aproximadamente

$$A = \frac{0.97}{\sqrt{15.05 + \sqrt[3]{0.98}}}$$

Tomando la función $f(x, y, z) = x/\sqrt{y + \sqrt[3]{z}}$ con $x = 1$, $y = 15$, $z = 1$, $dx = -0.03$, $dy = 0.05$, $dz = -0.02$. Tenemos

$$f(x + dx, y + dy, z + dz) = f(x, y, z) + df(x, y, z)$$

Como $f(x, y, z) = f(1, 15, 1) = \frac{1}{\sqrt{15+1}} = \frac{1}{4}$, sólo tenemos que calcular la diferencial de la función en $(1, 15, 1)$. Se tiene

$$\frac{\partial f}{\partial x} = \frac{1}{\sqrt{y + \sqrt[3]{z}}}, \qquad \frac{\partial f}{\partial y} = -\frac{x}{2}\left(y + \sqrt[3]{z}\right)^{-3/2},$$
$$\frac{\partial f}{\partial z} = -\frac{x}{2}\left(y + \sqrt[3]{z}\right)^{-3/2}\frac{1}{3}z^{-2/3}$$

Evaluado con $x = 1$, $y = 15$, $z = 1$ obtenemos

$$\frac{\partial f}{\partial x}(1, 15, 1) = \frac{1}{4}, \quad \frac{\partial f}{\partial y}(1, 15, 1) = -\frac{1}{128}, \quad \frac{\partial f}{\partial z} = -\frac{1}{384}$$

de modo que

$$df(1, 15, 1) = \frac{1}{4}(-0.03) - \frac{1}{128}(0.05) - \frac{1}{384}(-0.02) = -\frac{3.01}{384}$$

y así:

$$A \approx \frac{1}{4} - \frac{3.01}{384} = \frac{92.99}{384} = 0.242161$$

(el valor exacto es 0.2421726) ■

Ejercicios (Capítulo 2, Sección 11)

En los ejercicios 1–5, obtenga la diferencial de la función dada.

1. $f(x) = \operatorname{sen}^3 x^2$

2. $f(x, y) = x \tan y$

3. $f(x, y, z) = ax + by + cz + d$

4. $f(x, y, z, u) = \operatorname{sen} x + \cos y + \operatorname{arcsen} z + \arccos u$

5. $f(x, y, z, u, w) = xyz + xzw + yuw + zuw$

6. Calcule aproximadamente el incremento de la función $f(x, y) = x^y$ cuando la variable x pasa de $x_1 = 3$ a $x_2 = 3.1$, y la variable y pasa de $y_1 = 2$ a $y_2 = 1.9$.

7. Calcule aproximadamente el incremento de la función $f(x, y) = \frac{x^2-y^2}{3x+2y}$ cuando el punto (x, y) de su dominio pasa de $(2, 1)$ a $(2.05, 1.1)$.

8. Cuál es el incremento que sufre la función lineal $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ cuando su i-ésima variable x_i se incrementa en h_i unidades, $i = 1, 2, \ldots, n$?

En los ejercicios 9–14 aplique diferenciales para calcular aproximadamente el valor de la expresión dada.

9. $\sqrt{(2.13)(1.98)}$

10. $\ln((3.1)^2 - 8)$

11. $\ln(\sqrt{4.15} + \sqrt{9.08} - 4)$

12. $\arctan(\sqrt{0.2} + 0.98)$

13. $\sqrt[3.1]{(4.9)^2 + 2.1}$

14. $\operatorname{arcsen}\left(\sqrt[-1.9]{3.9 + (0.1)^{2.92}}\right)$

15. Calcule la longitud del segmento de la recta $x = x_0$, $y = y_0$ que se encuentra entre la superficie $z = x^2 + y^2$, y su plano tangente en el origen. (Nota: este problema lo tiene que resolver "a golpe de vista").

16. Calcule la longitud del segmento de la recta $x = 1.2$, $y = 0.95$, que se encuentra entre el paraboloide $z = x^2 + 5y^2$ y su plano tangente en el punto $(1, 1, 6)$.

17. Suponga que la superficie $z = f(x, y)$, gráfica de la función diferenciable f, tiene en el punto $\mathbf{p} = (x_0, y_0)$ el plano tangente $z = z_0 = f(x_0, y_0)$ y que en una bola B con centro en $\mathbf{p}$ y radio $r > 0$ la gráfica de la superficie se encuentra por debajo del plano tangente mencionado. Determine la longitud del segmento de recta $x = x_1$, $y = y_1$ (donde (x_1, y_1) es un punto de la bola B) que se encuentra entre la superficie $z = f(x, y)$ y el plano $z = z_0$.

2.12 Derivadas parciales de ordenes superiores

Consideremos la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$. Si esta función es diferenciable, sabemos que existen las derivadas parciales $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ en cualquier punto $(x, y) \in U$. Podemos entonces considerar las funciones

$$\frac{\partial f}{\partial x}: U \subseteq \mathbb{R}^2 \to \mathbb{R} \quad \text{y} \quad \frac{\partial f}{\partial y}: U \subseteq \mathbb{R}^2 \to \mathbb{R}$$

de modo que en cada punto $(x, y) \in U$ les asocia las derivadas parciales $\frac{\partial f}{\partial x}(x, y)$ y $\frac{\partial f}{\partial y}(x, y)$ respectivamente. Puede ocurrir que estas funciones sean a su vez lo suficientemente bien portadas en U como para que podamos obtener de ellas sus derivadas parciales

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

(en un punto dado de U). Diremos en tal caso que estas son las derivadas parciales de segundo orden de la función f y escribiremos

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$$
$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$$

NOTA: Usando la notación f_x y f_y para denotar a las derivadas parciales respecto de x y respecto de y de la función $f(x, y)$ respectivamente, las derivadas parciales de segundo orden de esta función se escriben $(f_x)_x = f_{xx}$, $(f_x)_y = f_{xy}$, $(f_y)_x = f_{yx}$, $(f_y)_y = f_{yy}$. Con la notación $D_1 f(x, y)$ y $D_2 f(x, y)$, las derivadas parciales de segundo orden se verían como $D_1(D_1 f(x, y) = D_{11} f(x, y)$, $D_2(D_1 f(x, y) = D_{12} f(x, y)$, $D_1(D_2 f(x, y)) = D_{21} f(x, y)$, $D_2(D_2 f(x, y)) = D_{22} f(x, y)$.

Suponiendo que las derivadas de 2° orden de $f(x, y)$ existan, podemos considerar ahora las funciones $\frac{\partial^2 f}{\partial x^2}: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $\frac{\partial^2 f}{\partial x \partial y}: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $\frac{\partial^2 f}{\partial y \partial x}: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $\frac{\partial^2 f}{\partial y^2}: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, que a cada punto $(x, y) \in U$ le asocian las derivadas parciales de segundo orden $\frac{\partial^2 f}{\partial x^2}(x, y)$, $\frac{\partial^2 f}{\partial x \partial y}(x, y)$, $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ y $\frac{\partial^2 f}{\partial y^2}(x, y)$ respectivamente. Y... la historia puede continuar. Si estas funciones son lo suficientemente bien portadas en U, podemos obtener de ellas sus derivadas parciales, las cuales serán las derivadas parciales de tercer orden de f en U (8 en total).

$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial^3 f}{\partial x^3} \qquad \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial^3 f}{\partial y \partial x^2}$$
$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial x^2 \partial y} \qquad \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial y \partial x \partial y}$$
$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x} \qquad \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial y^2 \partial x}$$
$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y^2}\right) = \frac{\partial^3 f}{\partial x \partial y^2} \qquad \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y^2}\right) = \frac{\partial^3 f}{\partial y^3}$$

etc.

Ejemplo 1. La función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = x^2 + y^2$ tiene por derivadas parciales a $\frac{\partial f}{\partial x} = 2x$, $\frac{\partial f}{\partial y} = 2y$. Las derivadas parciales de segundo orden de f son

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}(2x) = 2, \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}(2y) = 0$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}(2x) = 0, \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}(2y) = 2$$

■

Ejemplo 2. La función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = x^2 e^{x^2+y^2}$ tiene por derivadas parciales a

$$\frac{\partial f}{\partial x} = x^2 e^{x^2+y^2}(2x) + 2xe^{x^2+y^2} = 2e^{x^2+y^2}(x^3 + x)$$
$$\frac{\partial f}{\partial y} = x^2 e^{x^2+y^2}(2y) = 2x^2 ye^{x^2+y^2}$$

De estas nuevas funciones podemos obtener sus derivadas parciales y así tener las derivadas parciales de segundo orden de f

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(2xe^{x^2+y^2}(x^3 + x)\right)$$
$$= 2e^{x^2+y^2}(3x^2 + 1) + 4xe^{x^2+y^2}(x^3 + x) = 2e^{x^2+y^2}(2x^4 + 5x^2 + 1)$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}\left(2xe^{x^2+y^2}(x^3 + x)\right) = 4ye^{x^2+y^2}(x^3 + x)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(2x^2 ye^{x^2+y^2}\right)$$
$$= 2x^2 ye^{x^2+y^2}(2x) + 4xye^{x^2+y^2} = 4ye^{x^2+y^2}(x^3 + x)$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(2x^2 ye^{x^2+y^2}\right)$$
$$= 2x^2 ye^{x^2+y^2}(2y) + 2x^2 e^{x^2+y^2} = 2x^2 e^{x^2+y^2}(2y^2 + 1)$$ ■

En los dos ejemplos anteriores se observa que las derivadas parciales $\frac{\partial^2 f}{\partial x \partial y}$ y $\frac{\partial^2 f}{\partial y \partial x}$, llamadas *derivadas parciales cruzadas* (o mixtas) son iguales. Esta situación ocurre en muchas ocasiones (veremos el teorema correspondiente que da condiciones bajo las cuales esto ocurre), pero no es un hecho general, como muestra el ejemplo siguiente.

Ejemplo 3. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función

$$f(x, y) = \begin{cases} \dfrac{x^3 y^2 - xy^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

Si $(x, y) \neq (0, 0)$ las derivadas parciales de esta función son

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(\frac{x^3 y - xy^3}{x^2 + y^2}\right) = \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2}$$
$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(\frac{x^3 y - xy^3}{x^2 + y^2}\right) = \frac{x^5 - 4x^3 y^2 - xy^4}{(x^2 + y^2)^2}$$

Ahora bien, en el punto $(0, 0)$ las derivadas parciales de f son

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{\frac{h^3(0) - h(0)^3}{h^2 + (0)^2} - 0}{h} = 0$$
$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \to 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \to 0} \frac{\frac{(0)^3 h - 0h^3}{(0)^2 + h^2} - 0}{h} = 0$$

Calculemos $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ y $\frac{\partial^2 f}{\partial y \partial x}(0,0)$, usando directamente la definición de derivadas parciales

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x}(0,0) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = \lim_{h\to 0}\frac{\frac{\partial f}{\partial x}(0,h) - \frac{\partial f}{\partial x}(0,0)}{h}\\ &= \lim_{h\to 0}\frac{\frac{(0)^4 h + 4(0)^2 h^3 - h^5}{(0^2+h^2)^2} - 0}{h} = \lim_{h\to 0}(-1) = -1\\ \frac{\partial^2 f}{\partial x \partial y}(0,0) &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = \lim_{h\to 0}\frac{\frac{\partial f}{\partial y}(h,0) - \frac{\partial f}{\partial y}(0,0)}{h}\\ &= \lim_{h\to 0}\frac{\frac{h^5 - 4h^3(0) - h(0)^4}{(h^2+0^2)^2} - 0}{h} = \lim_{h\to 0}(1) = 1\end{aligned}$$

Por lo que

$$1 = \frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0) = -1$$ ■

El siguiente teorema nos dice cuándo debemos esperar que las derivadas parciales cruzadas de una función sean iguales.

Teorema 2.12.1 (Teorema de Schwarz) Sea $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el abierto U de $\mathbb{R}^2$. Si las derivadas $\frac{\partial^2 f}{\partial x \partial y}\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ y $\frac{\partial^2 f}{\partial y \partial x}\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ (existen y) son funciones continuas en U, entonces

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Demostración. (Opcional). Sea $(x, y) \in U$ un punto abierto de U. Para h y para k suficientemente pequeños (no nulos) los puntos $(x+h, y+k)$, $(x+h, y)$, $(x, y+k)$ siguen perteneciendo a U. Considere la expresión

$$M = f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x, y)$$

Sea $\varphi(x) = f(x, y+k) - f(x, y)$. Obsérvese entonces que

$$M = \varphi(x+h) - \varphi(x)$$

Aplicando el teorema del valor medio en la función φ en el intervalo $[x, x+h]$ obtenemos

$$M = \varphi(x+h) - \varphi(x) = \varphi'(\xi)h, \quad x < \xi < x+h$$

Obsérvese que

$$\varphi'(x) = \frac{\partial f}{\partial x}(x, y+k) - \frac{\partial f}{\partial x}(x, y)$$

de modo que

$$\varphi'(\xi) = \frac{\partial f}{\partial x}(\xi, y+k) - \frac{\partial f}{\partial x}(\xi, y)$$

y entonces M se ve

$$M = \varphi(x+h) - \varphi(x) = \varphi'(\xi)h = \left(\frac{\partial f}{\partial x}(\xi, y+k) - \frac{\partial f}{\partial x}(\xi, y)\right)h$$

Si consideramos ahora la función $\psi(y) = \frac{\partial f}{\partial x}(x, y)$, tenemos, aplicando el teorema del valor medio en el intervalo $[y, y+k]$, que

$$\psi(y+k) - \psi(y) = \psi'(\eta)h, \quad y < \eta < y+k$$

o sea

$$\frac{\partial f}{\partial x}(\xi, y+k) - \frac{\partial f}{\partial x}(\xi, y) = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)k$$

pues $\psi'(y)$ significa derivar respecto de y a $\frac{\partial f}{\partial x}$. Sustituyendo esta expresión en la obtenida anteriormente para M llegamos a

$$M = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)hk, \qquad \begin{array}{l} x < \xi < x+h \\ y < \eta < y+k \end{array}$$

Repitiendo la misma jugada, pero considerando primero la función $\bar{\varphi}(y) = f(x+h, y) - f(x, y)$ obtenemos

$$\begin{aligned} M = \bar{\varphi}(y+h) - \bar{\varphi}(y) = \bar{\varphi}'(\bar{\eta})k &= \left(\frac{\partial f}{\partial y}(x+h, \bar{\eta}) - \frac{\partial f}{\partial y}(x, \bar{\eta})\right)k \\ &= \frac{\partial^2 f}{\partial x \partial y}(\bar{\xi}, \bar{\eta})hk \end{aligned}$$

donde

$$x < \bar{\xi} < x+h, \quad y < \bar{\eta} < y+k$$

de modo que, igualando las dos expresiones para M y simplificando por hk llegamos a

$$\frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) = \frac{\partial^2 f}{\partial x \partial y}(\bar{\xi}, \bar{\eta})$$

Tomado límite cuando $h \to 0$ y $k \to 0$ y usando la continuidad asumida de las parciales mixtas, vemos que $\xi, \bar{\xi} \to x$, $\eta, \bar{\eta} \to y$, de modo que

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y)$$

como se quería. Q.E.D.

Muchas de las funciones que aparecen de manera natural en la práctica cumplen con la hipótesis del teorema de Schwarz, de modo que para ellas se tendrán solamente 3 derivadas parciales distintas de segundo orden. Más aún, en muchas funciones sus derivadas parciales cumplen también con la hipótesis del Teorema de Schwarz, por lo que

$$\frac{\partial^2}{\partial x \partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2}{\partial y \partial x}\left(\frac{\partial f}{\partial x}\right)$$

y

$$\frac{\partial^2}{\partial x \partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2}{\partial y \partial x}\left(\frac{\partial f}{\partial y}\right).$$

Es decir que $\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x^2}$ y $\frac{\partial^3 f}{\partial x \partial y^2} = \frac{\partial^3 f}{\partial y \partial x \partial y}$ y, como además (por ser iguales las parciales mixtas) se tiene

$$\frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x}$$

$$\frac{\partial^3 f}{\partial y \partial x \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial y^2 \partial x}$$

se ve que de las 8 derivadas parciales de tercer orden de la función f, se tienen sólo 4 derivadas distintas, a saber

1. $\frac{\partial^3 f}{\partial x^3}$
2. $\frac{\partial^3 f}{\partial x^2 \partial y}\left(= \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x^2}\right)$
3. $\frac{\partial^3 f}{\partial y^2 \partial x}\left(= \frac{\partial^3 f}{\partial y \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y^2}\right)$
4. $\frac{\partial^3 f}{\partial y^3}$

Y así entonces, para una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ que tenga todas sus derivadas parciales de todos los órdenes continuas en U, se tendrán sólo $(n + 1)$ derivadas distintas de orden n (de las 2^n que en principio existen). Estas son

$$\frac{\partial^n f}{\partial x^n}, \frac{\partial^n f}{\partial x^{n-1} \partial y}, \frac{\partial^n f}{\partial x^{n-2} \partial y^2}, \ldots, \frac{\partial^n f}{\partial x \partial y^{n-1}}, \frac{\partial^n f}{\partial y^n}$$

Ejemplo 3. Para una función polinomial $f: \mathbb{R}^2 \to \mathbb{R}$, todas sus derivadas parciales son también funciones polinomiales y por tanto son continuas. Entonces tendrá sólo $(n + 1)$ derivadas parciales distintas de orden n. Veamos esto con $n = 3$ y $f(x, y) = x^3 + 6x^2y^4 + 7xy^5 + 10x^3y$

Las derivadas de primer orden son

$$\frac{\partial f}{\partial x} = 3x^2 + 12xy^4 + 7y^5 + 30x^2y, \quad \frac{\partial f}{\partial y} = 24x^2y^3 + 35xy^4 + 10x^3$$

Las derivadas de segundo orden son

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(3x^2 + 12xy^4 + 7y^5 + 30x^2y\right) = 6x + 12y^4 + 60xy$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(24x^2y^3 + 35xy^4 + 10x^3\right) = 48xy^3 + 35y^4 + 30x^2$$

$$= \frac{\partial}{\partial y}(3x^2 + 12xy^4 + 7y^5 + 30x^2y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}(24x^2y^3 + 35xy^4 + 10x^3) = 72x^2y^2 + 140xy^3$$

Las derivadas de tercer orden son

$$\frac{\partial^3 f}{\partial x^3} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial}{\partial x}(6x + 12y^4 + 60xy) = 6 + 60y$$
$$\frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial x}(48xy^3 + 35y^4 + 30x^2) = 48y^3 + 60x$$
$$\frac{\partial^3 f}{\partial y^2 \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial}{\partial y}(48xy^3 + 35y^4 + 30x^2) = 144xy^2 + 140y^3$$
$$\frac{\partial^3 f}{\partial y^3} = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y^2}\right) = \frac{\partial}{\partial y}(72x^2y^2 + 140xy^3) = 144x^2y + 420xy^2$$

■

Ejemplo 4. Verifiquemos que la función $f(x, y) = \frac{x + y}{x^2 + y^2}$, satisface la ecuación

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

Las derivadas parciales de la función f son

$$\frac{\partial f}{\partial x} = \frac{y^2 - 2xy - x^2}{(x^2 + y^2)^2}$$
$$\frac{\partial^2 f}{\partial x^2} = \frac{2x^3 - 2y^3 + 6x^2y - 6xy^2}{(x^2 + y^2)^3}$$
$$\frac{\partial f}{\partial y} = \frac{x^2 - 2xy - y^2}{(x^2 + y^2)^2}$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{2y^3 - 2x^3 + 6xy^2 - 6xy^2}{(x^2 + y^2)^3}$$

de modo que

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{2x^3 - 2y^3 + 6x^2y - 6xy^2}{(x^2 + y^2)^3} + \frac{2y^3 - 2x^3 + 6xy^2 - 6xy^2}{(x^2 + y^2)^3} = 0$$

como se quería. ■

Por último mencionamos que para una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el abierto U de $\mathbb{R}^n$ que tenga sus derivadas parciales de segundo orden $\frac{\partial^2 f}{\partial x_i \partial x_j}$ continuas en U, entonces el Teorema de Schwarz establece que

$$\frac{\partial^2 f}{\partial x_i \partial y_j} = \frac{\partial^2 f}{\partial y_j \partial x_i}, \quad i, j = 1, \ldots, n$$

Este es un hecho que se deduce directamente del caso probado anteriormente: lo dejamos como un sencillo ejercicio para que el lector dé un argumento que establezca su validez.

Apéndice I Funciones de clase $\mathscr{C}^k$

Al considerar la propiedad de diferenciabilidad de una función $f: I \subseteq \mathbb{R} \to \mathbb{R}$, es natural preguntarse por las propiedades que presenta la función derivada $f': I \subseteq \mathbb{R} \to \mathbb{R}$. Por ejemplo: ¿esta función $f'(x)$ es continua?, ¿es derivable? La discusión que se presenta a este respecto con funciones de una sola variable se traslada sin gran dificultad, con las debidas adecuaciones en las definiciones —que hemos estudiado en este capítulo—, al caso de funciones de varias variables. Es por eso que presentamos al principio esta discusión en el caso de funciones reales de una variable real y, hacia el final del apéndice, extendemos los conceptos al caso de funciones de varias variables, objetivo de este capítulo.

Sea $f: I \subseteq \mathbb{R} \to \mathbb{R}$ una función diferenciable. Si la función derivada $f': I \subseteq \mathbb{R} \to \mathbb{R}$ es continua (en I), se dice que f es una *función de clase* $\mathscr{C}^1$. Si esta función f' es a su vez una función diferenciable, decimos que f es una *función dos veces diferenciable*. En tal caso consideramos la función $f'': I \subseteq \mathbb{R} \to \mathbb{R}$. Si ésta es continua, decimos que f es una *función de clase* $\mathscr{C}^2$. En general, decimos que la función f es una *función k veces diferenciable* si la función $f^{(k-1)}: I \subseteq \mathbb{R} \to \mathbb{R}$ es diferenciable. Si además ésta es continua, decimos que f es una *función de clase* $\mathscr{C}^k$. Una función f que es k veces diferenciable para todo $k \in \mathbb{N}$, se dice ser *infinitamente diferenciable*, o bien, *de clase* $\mathscr{C}^\infty$. Es claro entonces que se tienen las contenciones

$$\begin{pmatrix}\text{funciones}\\ \text{diferenciables}\end{pmatrix} \supset \begin{pmatrix}\text{funciones de}\\ \text{clase } \mathscr{C}^1\end{pmatrix} \supset \begin{pmatrix}\text{funciones dos}\\ \text{veces diferenciables}\end{pmatrix} \supset \begin{pmatrix}\text{funciones de}\\ \text{clase } \mathscr{C}^2\end{pmatrix} \supset \cdots \supset \begin{pmatrix}\text{funciones de}\\ \text{clase } \mathscr{C}^\infty\end{pmatrix}$$

Vemos que estas contenciones son estrictas. Por ejemplo, la función $f: \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) = \begin{cases} x^2 \operatorname{sen} \dfrac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

es diferenciable. De hecho

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} \frac{h^2 \operatorname{sen} \frac{1}{h}}{h} = \lim_{h \to 0} h \operatorname{sen} \frac{1}{h} = 0$$

y, para $x \neq 0$, se tiene

$$f'(x) = -\cos \frac{1}{x} + 2x \operatorname{sen} \frac{1}{x}$$

Es decir, la función derivada $f': \mathbb{R} \to \mathbb{R}$ es

$$f'(x) = \begin{cases} -\cos \dfrac{1}{x} + 2x \operatorname{sen} \dfrac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

la cual es una función discontinua (en $x = 0$, ¿por qué?). Entonces esta función, siendo diferenciable, *no es de clase* $\mathscr{C}^1$.

Por otra parte, la función $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^{5/3}$ es de clase $\mathscr{C}^1$, pues la función derivada $f': \mathbb{R} \to \mathbb{R}$, $f'(x) = \frac{5}{3}x^{2/3}$ es continua. Sin embargo, esta última función no es diferenciable, ya que su derivada (la segunda derivada de f), $f''(x) = \frac{10}{9}x^{-1/3}$, no existe en $x = 0$. Así pues, la

función $f(x) = x^{5/3}$ es de clase $\mathscr{C}^1$ pero *no es dos veces diferenciable*. Generalmente las funciones $f_n: \mathbb{R} \to \mathbb{R}$, $n \in \mathbb{N}$, dadas por

$$f_n(x) = x^n|x| = \begin{cases} x^{n+1} & \text{si } x \geq 0 \\ -x^{n+1} & \text{si } x < 0 \end{cases}$$

son diferenciables excepto en el caso $n = 0$. Su k-ésima derivada, $1 \leq k \leq n$, $f_n^{(k)}: \mathbb{R} \to \mathbb{R}$ igual a

$$f_n^{(k)}(x) = (n+1)(n)\dots(n-k+2)f_{n-k}(x)$$

de modo que en tanto $k < n$, las funciones $f_n^{(k)}(x)$ siguen siendo diferenciables. Sin embargo, con $k = n$ tenemos $f_n^{(n)}(x) = (n+1)!f_0(x) = (n+1)!|x|$, función que sólo es continua, pero no diferenciable (en $x = 0$). Así pues, las funciones $f_n(x) = x^n|x|$ son de clase $\mathscr{C}^n$, pero no son $(n+1)$ veces diferenciables. En particular, éstas son ejemplos de funciones de clase $\mathscr{C}^n$ que no son de clase $\mathscr{C}^{n+1}$.

En el caso de funciones de varias variables, la discusión anterior se presenta como sigue: se dice que la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U es de clase $\mathscr{C}^1$ (en U) si sus derivadas parciales $\frac{\partial f}{\partial x_i}: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, n$, (existen y) son funciones continuas en U.

Con base en el teorema 2.6.3, podemos concluir directamente el corolario siguiente.

Corolario (del Teorema 2.6.3) Si la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es de clase $\mathscr{C}^1$ entonces f es diferenciable. ■

Por supuesto, se tiene la misma situación que en el caso de funciones de una variable: el conjunto de funciones de clase $\mathscr{C}^1$ está *estrictamente* contenido en el conjunto de funciones diferenciables. Es decir, la afirmación recíproca del corolario anterior es falsa (ver ejercicio 28 de la sección 6).

Si además la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ tiene todas sus derivadas parciales de segundo orden continuas en U, entonces se dice que f es una función de clase $\mathscr{C}^2$ (en U), (nótese que no se definió "función dos veces diferenciable"; ver ejercicio 33 al final de la sección). En general, si en la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ sus derivadas parciales de k-ésimo orden son funciones continuas en U, entonces se dice que f es una función de clase $\mathscr{C}^k$ (en U). Si la función f tiene sus derivadas parciales de todos los órdenes continuas en U, se dice que f es una función de clase $\mathscr{C}^\infty$. Si por el contrario, la función f es solamente continua en U (digamos que sus derivadas parciales de primer orden no existen al menos en un punto de U), se dice que la función es de clase $\mathscr{C}^0$ (C cero).

Existen ejemplos de funciones $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ que muestran que las contenciones

$$\mathscr{C}^\infty \subset \dots \subset \mathscr{C}^2 \subset \mathscr{C}^1 \subset \mathscr{C}^0$$

son estrictas. Estos son, como era de esperarse, mucho más complicados que los que se dieron anteriormente para funciones de una sola variable.

Apéndice II El teorema de Euler sobre funciones homogéneas (versión general para funciones de dos variables).

En el apéndice de la sección 7 de este capítulo se estableció el resultado conocido como Teorema de Euler para funciones homogéneas, cuyo corolario dice que siendo $f: \mathbb{R}^n \to \mathbb{R}$ una función

diferenciable, homogénea de grado α, entonces

$$x_1 \frac{\partial f}{\partial x_1}(x) + x_2 \frac{\partial f}{\partial x_2}(x) + \cdots + x_n \frac{\partial f}{\partial x_n}(x) = \alpha f(x)$$

donde $x = (x_1, x_2, \ldots, x_n)$.

En el ejemplo 4 de ese apéndice, se consideró la función

$$f(x, y) = 8x^3 + 3x^2 y + 5y^3$$

la cual es homogénea de grado 3, con derivadas parciales

$$\frac{\partial f}{\partial x} = 24x^2 + 6xy, \qquad \frac{\partial f}{\partial y} = 3x^2 + 15y^2$$

y entonces

$$\begin{aligned} x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} &= x(24x^2 + 6xy) + y(3x^2 + 15y^2) \\ &= 24x^3 + 6x^2 y + 3x^2 y + 15y^3 \\ &= 24x^3 + 9x^2 y + 15y^3 \\ &= 3(8x^3 + 3x^2 y + 5y^3) = 3f(x, y) \end{aligned}$$

tal como tenía que ocurrir. Tomemos esta misma función para ejemplificar el resultado que queremos establecer en este apéndice. Las derivadas parciales de segundo orden de f son

$$\frac{\partial^2 f}{\partial x^2} = 48x + 6y, \qquad \frac{\partial^2 f}{\partial x \partial y} = 6x, \qquad \frac{\partial^2 f}{\partial y^2} = 30y$$

Se tiene que

$$\begin{aligned} x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} &= x^2(48x + 6y) + 2xy(6x) + y^2(30y) \\ &= 48x^3 + 18x^2 y + 30y^3 \\ &= (3)(2)(8x^3 + 3x^2 y + 5y^3) = (3)(3-1)f(x, y) \end{aligned}$$

Así, siendo f una función homogénea de grado 3, hemos verificado que se cumple la relación

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = (3)(3-1)f(x, y)$$

Deseándolo, podemos identificar el lado izquierdo de esta expresión como un binomio al cuadrado aplicado a la función f. De hecho, si definimos

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f = x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2}$$

para una función f suficientemente bien portada como para que sus parciales cruzadas sean iguales, hemos verificado que, siendo f homogénea de grado 3, se cumplen las relaciones

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 3f$$

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f = 3(3-1)f$$

Más aún, si vamos a las derivadas de tercer orden de la función f, se tiene

$$\frac{\partial^3 f}{\partial x^3} = 48, \quad \frac{\partial^3 f}{\partial x^2 \partial y} = 6, \quad \frac{\partial^3 f}{\partial x \partial y^2} = 0, \quad \frac{\partial^3 f}{\partial y^3} = 30$$

y entonces

$$\begin{aligned} x^3\frac{\partial^3 f}{\partial x^3} + 3x^2y\frac{\partial^3 f}{\partial x^2\partial y} + 3xy^2\frac{\partial^3 f}{\partial x\partial y^2} + y^3\frac{\partial^3 f}{\partial y^3} &= x^3(48) + 3x^2y(6) + 3xy^2(0) + y^3(30) \\ &= (3)(2)(1)(8x^3 + 3x^2y + 5y^3) \\ &= (3)(2)(1)f(x,y) \end{aligned}$$

Nuevamente, si definimos

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^3 f = x^3\frac{\partial^3 f}{\partial x^3} + 3x^2y\frac{\partial^3 f}{\partial x^2\partial y} + 3xy^2\frac{\partial^3 f}{\partial x\partial y^2} + y^3\frac{\partial^3 f}{\partial y^3}.$$

hemos probado que

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^3 f = (3)(3-1)(3-2)f$$

Este ejemplo nos pone en la mira la siguiente conjetura: si f es homogénea de grado n, y es lo suficientemente bien portada (desde el punto de vista de la diferenciabilidad; por ejemplo, que tenga derivadas parciales continuas de "muchos" órdenes), se cumple la relación

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^k f = n(n-1)\ldots(n-k+1)f$$

donde el lado izquierdo de esta expresión se define como

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^k f = \sum_{j=0}^{k}\binom{k}{j}x^{k-j}y^j\frac{\partial^k f}{\partial x^{k-j}\partial y^j}$$

La relación anterior es efectivamente cierta, y dar una demostración de ella es el objetivo de este apéndice. De hecho, el resultado vale para funciones de n variables: si $f\colon \mathbb{R}^n \to \mathbb{R}$ es una función definida en el conjunto abierto U de $\mathbb{R}^n$, homogénea de grado m, con "buen comportamiento diferenciable", entonces se tiene la relación

$$\boxed{\left(x_1\frac{\partial}{\partial x_1} + x_2\frac{\partial}{\partial x_2} + \cdots + x_n\frac{\partial}{\partial x_n}\right)^k f = m(m-1)(m-2)\cdots(m-k+1)f} \qquad \text{(e)}$$

para cualquier $k \in \mathbb{N}$. Las ideas que se expondrán para demostrar el caso $n = 2$ se pueden generalizar para probar la validez de la fórmula (e). Esto, sin embargo, nos obliga a pagar un precio muy alto en el aspecto de la notación, con el consiguiente peligro de perder las ideas centrales del argumento. Presentaremos entonces solamente la demostración del caso $n = 2$ y dejamos para el lector interesado que ajuste el argumento al caso general.

Establezcamos entonces formalmente el resultado que queremos probar. Como siempre, consideraremos funciones bien portadas desde el punto de vista diferenciable.

Teorema (De Euler sobre funciones homogéneas). Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable, homogénea de grado α. Entonces, para $k \in \mathbb{N}$ se tiene

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^k f(x, y) = \alpha(\alpha - 1)\ldots(\alpha - k + 1)f(x, y)$$

donde el primer miembro de esta expresión se interpreta como

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^k f(x, y) = \sum_{j=0}^{k} \binom{k}{j} x^{k-j} y^j \frac{\partial^k f}{\partial x^{k-j}\partial y^j}(x, y)$$ ■

Para la demostración de este teorema vamos a introducir un concepto que llamaremos "derivada euleriana" (o "derivada de Euler") de la función f, que denotaremos por $\mathcal{D}f(x, y)$ (o simplemente como $\mathcal{D}f$), y que definimos como

$$\mathcal{D}f = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) f$$

Esta derivada tiene un comportamiento lineal ante combinaciones lineales de funciones. Es decir, se tiene que $\mathcal{D}(f + cg) = \mathcal{D}f + c\mathcal{D}g$, donde $c \in \mathbb{R}$. Dejamos al lector que pruebe esta propiedad de la derivada euleriana (en el ejercicio 31 al final de esta sección se consideran otras propiedades).

Definimos la k-ésima derivada euleriana de f, denotada por $\mathcal{D}^k f$, como

$$\mathcal{D}^k f = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^k f$$

Por otra parte, definimos la acción de $\mathcal{D}$ dos veces sobre f, escrita $\mathcal{D}\mathcal{D}f$ como $\mathcal{D}\mathcal{D}f = \mathcal{D}(\mathcal{D}f)$. En general, $\mathcal{D}\mathcal{D}\ldots\mathcal{D}f$ significa $\mathcal{D}(\mathcal{D}(\ldots\mathcal{D}(\mathcal{D}f)\ldots))$. Debemos hacer notar que $\mathcal{D}^k f$ *no significa* la acción de $\mathcal{D}$ k veces sobre f. Es decir, $\mathcal{D}^k f$ no quiere decir $\mathcal{D}\mathcal{D}\mathcal{D}\ldots\mathcal{D}f$ (k veces $\mathcal{D}$). Por ejemplo, $\mathcal{D}^2 f$ no quiere decir $\mathcal{D}\mathcal{D}f$. En efecto, por una parte se tiene que $\mathcal{D}^2 f$ es

$$\mathcal{D}^2 f = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f = x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x\partial y} + y^2\frac{\partial^2 f}{\partial y^2}$$

y por otra parte

$$\begin{aligned}
\mathscr{D}\mathscr{D}f &= \mathscr{D}\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right)f = \mathscr{D}\left(x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}\right)\\
&= \left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right)\left(x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}\right)\\
&= x\frac{\partial}{\partial x}\left(x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}\right)+y\frac{\partial}{\partial y}\left(x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}\right)\\
&= x\left(x\frac{\partial^2 f}{\partial x^2}+\frac{\partial f}{\partial x}+\frac{\partial^2 f}{\partial x\partial y}\right)+y\left(x\frac{\partial^2 f}{\partial x\partial y}+y\frac{\partial^2 f}{\partial y^2}+\frac{\partial f}{\partial y}\right)\\
&= x^2\frac{\partial^2 f}{\partial x^2}+2xy\frac{\partial^2 f}{\partial x\partial y}+y^2\frac{\partial^2 f}{\partial y^2}+x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y} = \mathscr{D}^2 f+\mathscr{D}f
\end{aligned}$$

De hecho, vamos ahora a probar que

$$\mathscr{D}(\mathscr{D}^{k-1}f) = \mathscr{D}^k f+(k-1)\mathscr{D}^{k-1}f \qquad *$$

para $k \in \mathbb{N}$ (la hemos verificado para $k=2$). En efecto (por inducción sobre k), para $k=1$ es clara la relación (definiendo $\mathscr{D}^0 f$ como siendo f). Supongámosla válida para k y probémosla para $k+1$. Se tiene

$$\begin{aligned}
\mathscr{D}(\mathscr{D}^k f) &= x\frac{\partial}{\partial x}\mathscr{D}^k f+y\frac{\partial}{\partial y}\mathscr{D}^k f\\
&= x\frac{\partial}{\partial x}\sum_{j=0}^{k}\binom{k}{j}x^{k-j}y^j\frac{\partial^k f}{\partial x^{k-j}\partial y^j}+y\frac{\partial}{\partial y}\sum_{j=0}^{k}\binom{k}{j}x^{k-j}y^j\frac{\partial^k f}{\partial x^{k-j}\partial y^j}\\
&= x\sum_{j=0}^{k}\binom{k}{j}\left[(k-j)x^{k-j-1}y^j\frac{\partial^k f}{\partial x^{k-j}\partial y^j}+x^{k-j}y^j\frac{\partial^{k+1} f}{\partial x^{k-j+1}\partial y^j}\right]\\
&\quad+y\sum_{j=0}^{k}\binom{k}{j}\left[jx^{k-j}y^{j-1}\frac{\partial^k f}{\partial x^{k-j}\partial y^j}+x^{k-j}y^j\frac{\partial^{k+1} f}{\partial x^{k-j}\partial y^{j+1}}\right]\\
&= k\sum_{j=0}^{k}\binom{k}{j}x^{k-j}y^j\frac{\partial^k f}{\partial x^{k-j}\partial y^j}\\
&\quad+\sum_{j=0}^{k}\binom{k}{j}\left[x^{k-j+1}y^j\frac{\partial^{k+1} f}{\partial x^{k-j+1}\partial y^j}+x^{k-j}y^{j+1}\frac{\partial^{k+1} f}{\partial x^{k-j}\partial y^{j+1}}\right]\\
&= k\mathscr{D}^k f+\sum_{j=0}^{k}\binom{k}{j}x^{k-j+1}y^j\frac{\partial^{k+1} f}{\partial x^{k-j+1}\partial y^j}+\sum_{j=0}^{k}\binom{k}{j}x^{k-j}y^{j+1}\frac{\partial^{k+1} f}{\partial x^{k-j}\partial y^{j+1}}\\
&= k\mathscr{D}^k f+\sum_{j=0}^{k}\binom{k}{j}x^{k-j+1}y^j\frac{\partial^{k+1} f}{\partial x^{k-j+1}\partial y^j}+\sum_{j=1}^{k+1}\binom{k}{j-1}x^{k-j+1}y^j\frac{\partial^{k+1} f}{\partial x^{k-j+1}\partial y^j}\\
&= k\mathscr{D}^k f+x^{k+1}\frac{\partial^{k+1} f}{\partial x^{k+1}}+\sum_{j=1}^{k}\binom{k}{j}x^{k-j+1}y^j\frac{\partial^{k+1} f}{\partial x^{k-j+1}\partial y^j}
\end{aligned}$$

$$+\sum_{j=1}^{k}\binom{k}{j-1}x^{k-j+1}y^j\frac{\partial^{k+1}f}{\partial x^{k-j+1}\partial y^j}+y^{k+1}\frac{\partial^{k+1}f}{\partial y^{k+1}}$$

$$=k\mathscr{D}^k f+x^{k+1}\frac{\partial^{k+1}f}{\partial x^{k+1}}+y^{k+1}\frac{\partial^{k+1}f}{\partial y^{k+1}}$$

$$+\sum_{j=1}^{k}\left[\binom{k}{j}+\binom{k}{j-1}\right]x^{k-j+1}y^j\frac{\partial^{k+1}f}{\partial x^{k-j+1}\partial y^j}$$

$$=k\mathscr{D}^k f+x^{k+1}\frac{\partial^{k+1}f}{\partial x^{k+1}}+y^{k+1}\frac{\partial^{k+1}f}{\partial y^{k+1}}+\sum_{j=1}^{k}\binom{k+1}{j}x^{k-j+1}y^j\frac{\partial^{k+1}f}{\partial x^{k-j+1}\partial y^j}$$

$$=k\mathscr{D}^k f+\sum_{j=0}^{k+1}\binom{k+1}{j}x^{k-j+1}y^j\frac{\partial^{k+1}f}{\partial x^{k-j+1}\partial y^j}=k\mathscr{D}^k f+\mathscr{D}^{k+1}f$$

Así hemos probado que $\mathscr{D}(\mathscr{D}^k f)=k\mathscr{D}^k f+\mathscr{D}^{k+1}f$, que es la misma expresión (*) con $k+1$. Ya estamos ahora en posibilidades de probar nuestro teorema. Se quiere demostrar entonces que siendo f una función homogénea de grado α se cumple que

$$\mathscr{D}^k f=\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right)^k f=\alpha(\alpha-1)\cdots(\alpha-k+1)f$$

para cualquier entero positivo k. La prueba se hará por inducción sobre k. Para $k=1$ ya se ha demostrado el teorema (en el apéndice de la sección 7). Supongamos válida la relación anterior para k y probémosla para $k+1$. Se tiene, usando la fórmula (*) previamente demostrada (con $k+1$)

$$\mathscr{D}^{k+1}f=\mathscr{D}(\mathscr{D}^k f)-k\mathscr{D}^k f\ \overset{\text{hipótesis de inducción}}{=\!=\!\longrightarrow}$$

$$=\mathscr{D}\Big(\alpha(\alpha-1)\ldots(\alpha-k+1)f\Big)-k\alpha(\alpha-1)\ldots(\alpha-k+1)f$$

$$\overset{\text{linealidad de}\ \mathscr{D}}{=\!=\!\longrightarrow}$$

$$=\alpha(\alpha-1)\ldots(\alpha-k+1)\mathscr{D}f-k\alpha(\alpha-1)\ldots(\alpha-k+1)f$$

$$\underset{\mathscr{D}f=\alpha f}{\overset{\text{caso } k=1}{=\!=\!\longrightarrow}}$$

$$=\alpha(\alpha-1)\ldots(\alpha-k+1)\alpha f-k\alpha(\alpha-1)\ldots(\alpha-k+1)f$$

$$=\alpha(\alpha-1)\ldots(\alpha-k+1)(\alpha-k)f$$

que es la fórmula que el teorema establece con $k+1$.

Ejercicios (Capítulo 2, Sección 2.12)

En los ejercicios 1–5 calcule las derivadas parciales de segundo orden de la función dada.

1. $f(x,y)=x\operatorname{sen}y+y\operatorname{sen}x$

2. $f(x, y) = x^y$

3. $f(x, y) = \arctan(xy)$

4. $f(x, y) = \ln(x + y)$

5. $f(x, y) = (\operatorname{sen} x)(\operatorname{sen} y)$

6. Calcule las derivadas parciales de segundo orden de la función lineal $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$.

7. Sean $f_1, f_2, \ldots, f_n: \mathbb{R} \to \mathbb{R}$ funciones de clase $\mathscr{C}^2$. Calcule las derivadas parciales de segundo orden de la función $F: \mathbb{R}^n \to \mathbb{R}$ dada por

   a. $F(x_1, x_2, \cdots, x_n) = f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$
   b. $F(x_1, x_2, \cdots, x_n) = f_1^2(x_1) + f_2^2(x_2) + \cdots + f_n^2(x_n)$
   c. $F(x_1, x_2, \cdots, x_n) = f_1(x_1) + f_2^2(x_2) + \cdots + f_n^n(x_n)$
   d. $F(x_1, x_2, \cdots, x_n) = f_1(x_1)f_2(x_2)\cdots f_n(x_n)$
   e. $F(x_1, x_2, \cdots, x_n) = f_1^2(x_1)f_2^2(x_2)\ldots f_n^2(x_n)$
   f. $F(x_1, x_2, \cdots, x_n) = f_1(x_1)f_2^2(x_2)\ldots f_n^n(x_n)$

8. Sea $f: \mathbb{R}^n \to \mathbb{R}$ la función $f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n-1} x_ix_{i+1}$. Calcule las derivadas parciales de segundo orden de f.

Demuestre que cada una de las funciones dadas en los ejercicios 9–14 satisface la ecuación $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$ (llamada "ecuación de Laplace")

9. $f(x, y) = x^3 - 3xy^2$

10. $f(x, y) = e^x(\cos y + \operatorname{sen} y)$

11. $f(x, y) = e^x(x\cos y - y\cos y)$

12. $f(x, y) = e^{x^2-y^2}(\cos 2xy + \operatorname{sen} 2xy)$

13. $f(x, y) = \ln \frac{1}{\sqrt{x^2+y^2}}$

14. $f(x, y) = e^{x^2-y^2}(2xy\cos 2xy + (x^2 - y^2)\operatorname{sen} 2xy)$

15. Constate que la función $z = \operatorname{sen}(x^2 + y^2)$ satisface la ecuación

$$y\frac{\partial^2 z}{\partial x^2} - x\frac{\partial^2 z}{\partial y\partial x} - \frac{\partial z}{\partial y} = 0$$

16. Constate que la función $u = (x - at)^2 + (x + at)^3$ satisface la ecuación

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2}$$

(llamada "ecuación de calor").

17. En el ejercicio 62 de la sección 4 se estableció que si g es una función real de variable real, diferenciable, y $f(x, y)$ es una función con derivadas parciales, entonces la composición $g \circ f$ tiene derivadas parciales y éstas son

$$\frac{\partial}{\partial x}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial f}{\partial x}(\mathbf{p}), \quad \frac{\partial}{\partial y}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial f}{\partial y}(\mathbf{p})$$

donde $\mathbf{p} = (x, y)$ es un punto del dominio de f tal que $f(x, y)$ está en el dominio de g. Suponga ahora que tanto g como f tienen derivadas de segundo orden. Demuestre que

$$\frac{\partial^2}{\partial x^2}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial^2 f}{\partial x^2}(\mathbf{p}) + g''(f(\mathbf{p}))\left(\frac{\partial f}{\partial x}(\mathbf{p})\right)^2$$

$$\frac{\partial^2}{\partial y^2}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial^2 f}{\partial y^2}(\mathbf{p}) + g''(f(\mathbf{p}))\left(\frac{\partial f}{\partial y}(\mathbf{p})\right)^2$$

$$\frac{\partial^2}{\partial x \partial y}(g \circ f)(\mathbf{p}) = g'(f(\mathbf{p}))\frac{\partial^2 f}{\partial x \partial y}(\mathbf{p}) + g''(f(\mathbf{p}))\frac{\partial f}{\partial x}(\mathbf{p})\frac{\partial f}{\partial y}(\mathbf{p})$$

En los ejercicios 18–21, g es una función real de variable real, de clase $\mathscr{C}^2$. Tome el resultado del ejercicio anterior para calcular las derivadas de segundo orden de las funciones $F(x, y)$ indicadas.

18. $F(x, y) = g(2x + 3y)$

19. $F(x, y) = xg(xy)$

20. $F(x, y) = (x^2 + y^2)g(x^2 + y^2)$

21. $F(x, y) = g(g(x) + g(y))$

22. Sea $z = g(x^2 + y^2)$, donde g es una función real de variable real, dos veces derivable. Demuestre que

$$y\frac{\partial^2 z}{\partial x^2} - x\frac{\partial^2 z}{\partial y \partial x} - \frac{\partial z}{\partial y} = 0$$

Nótese que el ejercicio 15 es un caso particular de este ejercicio.

23. Sea $u = \phi(x - at) + \psi(x + at)$, donde ϕ y ψ son dos funciones reales de variable real, dos veces derivables. Demuestre que esta función $u = f(x, t)$ es solución de la ecuación de calor

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2}$$

Obsérvese que el ejercicio 16 es un caso particular de este ejercicio.

24. Sea $z = x\phi(x + y) + y\psi(x - y)$, donde ϕ y ψ son funciones reales de variable real, dos veces derivables. Demuestre que

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 2\frac{\partial^2 f}{\partial x \partial y}$$

25. Demuestre que la función del ejemplo 3, $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} \dfrac{x^3y - xy^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

no satisface las hipótesis del teorema de Schwarz.

26. En el apéndice de la sección 7 se estableció el teorema de Euler sobre funciones homogéneas: si $f(x, y)$ es homogénea de grado n, entonces

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = nf$$

Más adelante probaremos que si una función $f(x, y)$ satisface la relación anterior, entonces f debe ser homogénea de grado n (ejemplo 10 de la sección 2 del capítulo 3). Se tiene entonces que la relación anterior se satisface *si y sólo si* f es una función homogénea de grado n. En el apéndice II de esta sección, probamos que siendo f una función homogénea de grado n, entonces también se cumplía que

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = n(n-1)f$$

¿Será esta condición también suficiente para asegurar que f es homogénea de grado n? La respuesta es NO. Demuestre que la función $f(x, y) = xy + \frac{y}{x^2}$ satisface la relación anterior con $n = 2$, a pesar de que tal función *no es homogénea.*

27. Demuestre que si $\phi(x, y)$ es una función homogénea de grado n y $\psi(x, y)$ es una función homogénea de grado $1 - n$, entonces la función $f(x, y) = \phi(x, y) + \psi(x, y)$ (que en general no es homogénea) satisface la relación

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = n(n-1)f$$

Vuelva al ejercicio anterior y explique lo ahí demostrado a la luz de este ejercicio.

28. Tome el resultado del ejercicio anterior para demostrar que si ϕ y ψ son funciones homogéneas de grado cero, entonces la función $f(x, y) = \phi(x, y) + x\psi(x, y)$ satisface la relación

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x \partial y} + y^2\frac{\partial^2 f}{\partial y^2} = 0$$

29. Sea $z = f(x, y)$ una función homogénea de grado α. En el ejercicio 28 de la sección 7 se probó que las derivadas parciales de f, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ son funciones homogéneas de grado $\alpha - 1$. Use el teorema de Euler sobre funciones homogéneas aplicado a estas derivadas parciales para concluir que

$$x\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) + y\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = (\alpha - 1)\frac{\partial f}{\partial x}$$

$$x\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) + y\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = (\alpha - 1)\frac{\partial f}{\partial y}$$

Multiplique por x la primera de estas expresiones y por y la segunda, y luego sume las expresiones resultantes para obtener que

$$x^2\frac{\partial^2 f}{\partial x^2}+2xy\frac{\partial^2 f}{\partial x\partial y}+y^2\frac{\partial^2 f}{\partial y^2}=(\alpha-1)\left(x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}\right)=\alpha(\alpha-1)f$$

Esta es otra demostración del teorema de Euler sobre funciones homogéneas (caso $k=2$) estudiado en el apéndice II.

30. Considere la función $f\colon \mathbb{R}^2\to\mathbb{R}$ homogénea de grado α, suficientemente diferenciable. Demuestre que para $k\in\mathbb{N}$, se tiene

$$\mathcal{D}^k f = x\mathcal{D}^{k-1}\left(\frac{\partial f}{\partial x}\right)+y\mathcal{D}^{k-1}\left(\frac{\partial f}{\partial y}\right)$$

donde $\mathcal{D}$ es la derivada de Euler definida en el apéndice II. Utilice este hecho para una nueva demostración del teorema de Euler sobre funciones homogéneas.

31. Sean $f, g\colon U\subseteq\mathbb{R}^2\to\mathbb{R}$ dos funciones diferenciables. Demuestre las siguientes propiedades de la derivada de Euler:

a. $\mathcal{D}(f+cg)=\mathcal{D}f+c\mathcal{D}g,\ c\in\mathbb{R}$

b. $\mathcal{D}(fg)=f\mathcal{D}g+g\mathcal{D}f$

c. $\mathcal{D}\left(\frac{f}{g}\right)=\frac{g\mathcal{D}f-f\mathcal{D}g}{g^2}$, (para las $(x,y)\in U$ tales que $g(x,y)\neq 0$)

32. Demuestre que las funciones $f_n\colon\mathbb{R}\to\mathbb{R}$, $n\in\mathbb{N}$, dadas por

$$\begin{cases} x^{2n}\operatorname{sen}\dfrac{1}{x} & \text{si } x\neq 0\\ 0 & \text{si } x=0\end{cases}$$

son n veces diferenciables, pero no son de clase $\mathscr{C}^n$.

33. Para funciones de varias variables no se estableció el concepto de "funciones k veces diferenciables", saltando, desde la clase $\mathscr{C}^{k-1}$ hasta la clase $\mathscr{C}^k$. ¿Sugiere el lector alguna causa por la cual no se incluyó este concepto?, ¿qué sinificaría decir que la función $f\colon\mathbb{R}^2\to\mathbb{R}$ es dos veces diferenciable?

34. (Continuación del ejercicio 31 de la sección 6: un breve curso "hágalo usted mismo" de funciones de variable compleja. parte (ii): funciones armónicas).

q. Un resultado clásico del análisis complejo, que es consecuencia del teorema de Cauchy que estudiaremos en la cuarta parte de este problema (sección 4 del capítulo 7), en el capítulo 7, establece que si la función $f(z)$ es holomorfa (es decir, existe $f'(z)$ en todo z de su dominio), entonces esta función tiene derivadas de todos los órdenes. Es decir, la sola existencia de la primera derivada garantiza la existencia de las derivadas de todos los órdenes. En particular, (verifique que) si $f\colon U\subseteq\mathbb{C}\to\mathbb{C}$ es holomorfa, las funciones $u, v\colon U\subseteq\mathbb{R}^2\to\mathbb{R}$ son de clase $\mathscr{C}^2$. Use este hecho, junto con las ecuaciones de Cauchy-Riemann, y el teorema de Schwarz estudiado en esta sección para demostrar siendo $f=u+iv$ holomorfa, las funciones u y v satisfacen la ecuación de Laplace

$$\frac{\partial^2 F}{\partial x^2}+\frac{\partial^2 F}{\partial y^2}=0$$

(ver ejercicios 9–14). A una función $F(x, y)$ que satisface esta ecuación se le llama *función armónica*. Así pues, siendo $f = u + iv$ holomorfa, las funciones u y v son armónicas.

r. Considere la función holomorfa $f(z) = z^3$. Identifique sus partes real e imaginaria. Vea con nuevos ojos el ejercicio 9 de esta sección. (Este ejercicio continuará en el ejercicio 18 sección 3, del capítulo 7.)

Capítulo 3

Funciones compuestas, inversas e implícitas

En este capítulo continuaremos con el estudio del aspecto diferencial de las funciones de varias variables, considerando ahora funciones cuya descripción es, o más complicada que las del capítulo anterior (como serán las funciones compuestas), o bien, cuya descripción no es "explícita", como lo han sido las funciones estudiadas hasta este momento (las funciones implícitas).

Recordemos rápidamente el tipo de resultados que, en estos temas, aparecían en el estudio de funciones de una variable.

Si tenemos las funciones $g: I \subseteq \mathbb{R} \to \mathbb{R}$, $f: J \subseteq \mathbb{R} \to \mathbb{R}$ (tales que $g(I) \subseteq J$), podíamos formar la composición $f \circ g: I \subseteq \mathbb{R} \to \mathbb{R}$, definida como $(f \circ g)(x) = f(g(x))$. El resultado más importante de esta parte del curso (la regla de la cadena) nos dice que si g es diferenciable en un punto $x_0 \in I$ y f es diferenciable en $g(x_0) \in J$, entonces la composición $f \circ g$ es diferenciable en x_0, es decir, que $(f \circ g)'(x)$ existe, y su valor no es más que el producto de las derivadas de las funciones f y g. Más bien

$$(f \circ g)'(x) = f'(g(x))g'(x)$$

o bien, con la notación de Leibniz, si $y = f(u)$ y $u = g(x)$ entonces

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

(la ventaja de esta noción es obvia: sugiere a la vista que se cancela du y queda una igualdad trivial. Obsérvese sin embargo, que en ella no se hacen explícitos los puntos donde se calculan las derivadas).

En el caso de funciones de varias variables, la situación se hará (en principio) más complicada. Téngase en cuenta que, para empezar, no podemos hablar de composición de dos funciones $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$, pues la operación de composición entre funciones solamente se puede realizar cuando el rango de una de ellas está contenido en el dominio de la otra (y en nuestro caso ambas funciones f y g tienen rango en $\mathbb{R}$ y dominio en $\mathbb{R}^n$). Nuestro propósito en este capítulo será dejar claro el proceso de composición de funciones de varias variables y luego estudiar "la regla de la cadena" en el caso general.

En el caso de funciones inversas en una variable teníamos que si la función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ es inyectiva, con rango $J \subseteq \mathbb{R}$, se da la función $f^{-1}: J \subseteq \mathbb{R} \to \mathbb{R}$, llamada *inversa* de f, tal que $f^{-1}(f(x)) = x$, $x \in I$ y $f(f^{-1}(y)) = y$, $y \in J$. Sabiendo que f es una función diferenciable tal

que $f'(x_0) \neq 0$, $x_0 \in I$, entonces podíamos, al menos localmente, invertir f y obtener f^{-1} también diferenciable, cuya derivada se encontraba como

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}$$

o con la notación de Leibniz, si $y = f(x)$, entonces

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$$

donde $x = f^{-1}(y)$.

En el caso de funciones de varias variables esta situación también se complica, pues se tienen que considerar funciones más generales con condiciones más generales (piense: ¿tiene sentido hablar de "inversa" de una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$?). Este estudio nos llevará a uno de los célebres teoremas del análisis, llamado Teorema de la Función Inversa, el cual estudiaremos en la sección 6.

Por último, en el cálculo de funciones de una variable, se consideraron expresiones del tipo $F(x, y) = 0$, preguntándonos si podíamos despejar a y en términos de x, y dejar establecida una función del tipo $y = f(x)$ (por ejemplo si $F(x, y) = x^2 + y^2 - 1 = 0$ no se puede obtener de ella una función $y = f(x)$; y de $F(x, y) = x^2 - y = 0$ sí se puede obtener la función $y = f(x) = x^2$). Cuando de $F(x, y) = 0$ podemos despejar la variable y, y escribir $y = f(x)$, decimos que ésta última función está dada implícitamente en $F(x, y) = 0$. En realidad, el tema de las funciones implícitas no se puede abordar plenamente en el primer curso de cálculo, pues para entender bien las cuestiones que en él se tratan, se tiene que acudir a resultados más elaborados (como los de un curso de cálculo en $\mathbb{R}^n$). Con la herramienta que hemos desarrollado hasta este momento estaremos en condiciones de llegar a entender bien la problemática en torno a las funciones implícitas, estableciendo, en la sección 5, otro de los célebres teoremas del análisis: el Teorema de la Función Implícita.

Comencemos pues por estudiar cómo se hace la composición de funciones de varias variables.

3.1 Composición de funciones

Componer funciones significa "sustituir una función en otra". Una manera sencilla de entender esto es la siguiente: si tenemos la función $y = f(x)$ que establece la dependencia de y y x, y la función $x = g(u)$, que establece la dependencia entre x y u, podemos sustituir esta última en la primera y obtener $y = f(g(u))$. A la función así obtenida (que manda u a y) se le llama composición de f con g y se denota por $f \circ g$. En un esquema se ve como en la figura 1.

Obsérvese entonces que para obtener la función compuesta "se ha sustituido la variable x de la función f por la función g que conecta a esta variable con otra, a saber, u".

Pasemos al caso de funciones de dos variables $z = f(x, y)$. Siguiendo con la misma idea, para "componer" esta función tendremos que sustituir *las dos variables* x y y por *dos funciones*, digamos g_1 y g_2 que conecten a éstas con *otras* variables, digamos otras dos variables u y v. Así, si consideramos las funciones

$$x = g_1(u, v),\ y = g_2(u, v)$$

podemos sustituir éstas en la función f y obtener *la función compuesta*

$$y = f(g_1(u, v), g_2(u, v))$$

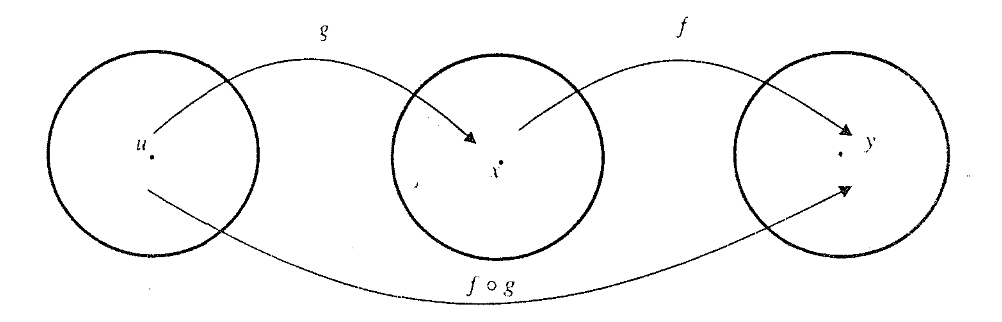


Figura 1. Composición de funciones.

En esta perspectiva, una función f de *dos* variables se compone de *dos* funciones que conectan cada una de estas variables con otras variables distintas.

Ejemplo 1. Considere la función $f(x, y) = x^2y + \text{sen}(xy)$. Si $x = g_1(u, v) = u^2 + v^2$, $y = g_2(u, v) = uv^3$, entonces la función compuesta de f con g_1 y g_2 es

$$\begin{aligned} f(g_1(u, v), g_2(u, v)) &= (g_1(u, v))^2 g_2(u, v) + \text{sen}(g_1(u, v)g_2(u, v)) \\ &= (u^2 + v^2)^2(uv^3) + \text{sen}((u^2 + v^2)(uv^3)) \end{aligned}$$

Nótese que ésta es una función *distinta* de f, g_1 y g_2; aquélla es la *composición* de f con g_1 y g_2. ■

Ejemplo 2. Sea $f(x, y, z) = 2x^3y^2z + \cos^2(x + y + z)$. Sean también las funciones

$$x = g_1(t) = e^t, \quad y = g_2(t) = \cos t, \quad z = g_3(t) = t^2$$

Entoces la composición de f con g_1, g_2 y g_3 es

$$\begin{aligned} f(g_1(t), g_2(t), g_3(t)) &= 2(g_1(t))^3(g_2(t))^2 g_3(t) + \cos^2(g_1(t) + g_2(t) + g_3(t)) \\ &= 2e^{3t}(\cos^2 t)t^2 + \cos^2(e^t + \cos t + t^2). \end{aligned}$$ ■

Una mejor manera de contemplar el proceso de composición de funciones es pensando en funciones de $\mathbb{R}^m$ en $\mathbb{R}^n$: la pareja de funciones $x = g_1(u, v)$, $y = g_2(u, v)$ las podemos ver como *una* función g que va de $\mathbb{R}^2$ en $\mathbb{R}^2$ de la siguiente manera. A cada punto $(u, v) \in \mathbb{R}^2$ la función g le asocia el punto $g(u, v) \in \mathbb{R}^2$ cuyas coordenadas serán $(x, y) = (g_1(u, v), g_2(u, v))$. (Figura 2)

A las funciones g_1 y g_2 se les llama (por razones obvias) las *funciones coordenadas* de la función g. Escribimos entonces $g: \mathbb{R}^2 \to \mathbb{R}^2$, $g(u, v) = (g_1(u, v), g_2(u, v))$. Por ejemplo, si $g_1(u, v) = u + v$, $g_2(u, v) = u^2 - v^2$, entonces la función $g: \mathbb{R}^2 \to \mathbb{R}^2$ es $g(u, v) = (u + v, u^2 - v^2)$. Si tuviéramos 3 funciones, cada una de ellas de una variable (como el ejemplo 2 anterior), digamos $x = g_1(t)$, $y = g_2(t)$, $z = g_3(t)$, podríamos juntar estas tres funciones en una sola función $g: \mathbb{R} \to \mathbb{R}^3$, de modo que $g(t) = (g_1(t), g_2(t), g_3(t))$. Así, las funciones $x = e^t$, $y = \cos t$, $z = t^2$ del ejemplo 2 se pueden ver como las funciones coordenadas de la función $g(t) = (e^t, \cos t, t^2)$.

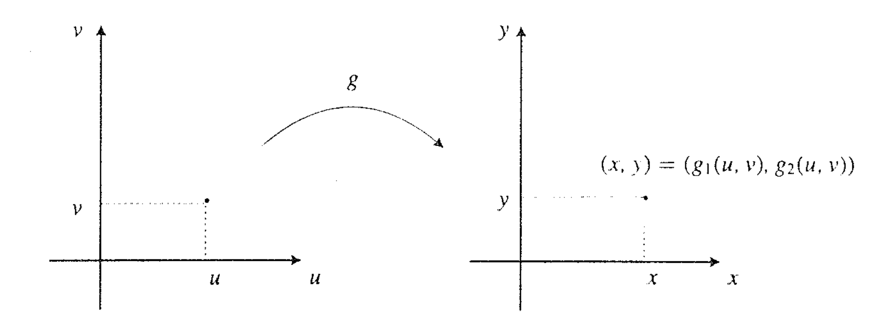


Figura 2. El punto (x, y) como imagen de $g_1(u, v)$ y $g_2(u, v)$.

En general, los conjuntos de n funciones reales, cada una de ellas dependiendo de m variables, digamos

$$\begin{aligned} x_1 &= g_1(u_1, u_2, \ldots, u_m) \\ x_2 &= g_2(u_1, u_2, \ldots, u_m) \\ &\vdots \\ x_n &= g_n(u_1, u_2, \ldots, u_m) \end{aligned}$$

se puede ver como *una* función $g: \mathbb{R}^m \to \mathbb{R}^n$ que toma al punto $(u_1, u_2, \ldots, u_m) \in \mathbb{R}^m$ y le asocia el punto $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, para la cual las funciones dadas son sus funciones coordenadas. Las funciones $g: U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ definidas en el conjunto abierto U de $\mathbb{R}^m$ y tomando valores en el espacio $\mathbb{R}^n$ son las funciones más generales que podemos encontrar en nuestro curso de cálculo. El caso $n = m = 1$, es el de las funciones reales de una variable real que se estudiaron en el primer curso. Si $n = 1$ y m es arbitrario, se tienen las funciones reales de m variables que ya estudiamos en el capítulo anterior. En el caso general, si el codominio es el espacio $\mathbb{R}^n$, debemos ver estas funciones como constituidas por n funciones reales de m variables (sus funciones coordenadas). En este punto se ve la importancia de estudiar primero las funciones de m variables $\tilde{g}: U \subseteq \mathbb{R}^m \to \mathbb{R}$, pues son los eslabones que componen a funciones más generales $g: U \subseteq \mathbb{R}^m \to \mathbb{R}^n$.

Ejemplo 3. La función $g: \mathbb{R}^2 \to \mathbb{R}^3$, dada por

$$g(u, v) = (u^2 - v^2, u + v, u - v)$$

tiene por funciones coordenadas a $g_1, g_2, g_3: \mathbb{R}^2 \to \mathbb{R}$, $g_1(u, v) = u^2 - v^2$, $g_2(u, v) = u + v$, $g_3(u, v) = u - v$. ■

Con esta perspectiva podemos ver mejor el proceso de composición de funciones de varias variables como una sustitución de *una función en otra* (es decir, como un proceso entre *dos* funciones).

Supongamos que tenemos la función $f: \mathbb{R}^n \to \mathbb{R}$ definida en el espacio $\mathbb{R}^n$. Si consideramos la función $g: \mathbb{R}^m \to \mathbb{R}^n$, cuyo codominio $\mathbb{R}^n$ es el dominio de f, podemos formar la composición $f \circ g: \mathbb{R}^m \to \mathbb{R}$, definida como $(f \circ g)(\mathbf{u}) = f(g(\mathbf{u}))$. (Figura 3)

Generalmente, si tenemos la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto U de $\mathbb{R}^n$ y la función $g: V \subseteq \mathbb{R}^m \to \mathbb{R}^n$ definida en el conjunto V de $\mathbb{R}^m$, cuyo rango está contenido en U

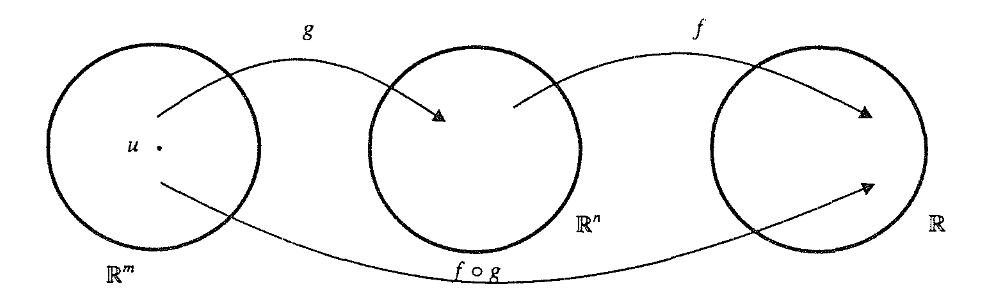


Figura 3. Composición de funciones.

(i.e. tal que $g(V) \subseteq U$), entonces podemos formar la composición $f \circ g: V \subseteq \mathbb{R}^m \to \mathbb{R}$, como $(f \circ g)(\mathbf{v}) = f(g(\mathbf{v}))$, $\mathbf{v} \in V$. (Figura 4).

Ejemplo 4. Considere la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ del ejemplo 1, $f(x, y) = x^2y + \operatorname{sen}(xy)$. Si componemos esta función con la función $g: \mathbb{R}^2 \to \mathbb{R}^2$, $g(u, v) = (u^2 + v^2, uv^3)$, se obtiene $f \circ g: \mathbb{R}^2 \to \mathbb{R}$, $(f \circ g)(u, v) = f(g(u, v)) = f(u^2 + v^2, uv^3) = (u^2 + v^2)^2(uv^3) + \operatorname{sen}((u^2 + v^2)uv^3)$, como se obtuvo en el ejemplo 1 componiendo f con las tres funciones coordenadas de g. De la misma manera, podemos ver el ejemplo 2 como la composición de la función $f: \mathbb{R}^3 \to \mathbb{R}$, dada por $f(x, y, z) = 2x^3y^2z + \cos^2(x + y + z)$, con la función $g: \mathbb{R} \to \mathbb{R}^3$ dada por $g(t) = (e^t, \cos t, t^2)$, obteniéndose la función compuesta $f \circ g: \mathbb{R} \to \mathbb{R}$, dada por $(f \circ g)(t) = f(g(t)) = f(e^t, \cos t, t^2) = 2e^{3t}(\cos^2 t)t^2 + \cos^2(e^t + \cos t + t^2)$. ■

Ejemplo 5. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = 3x^2 + 5xy$ y sea $g: \mathbb{R}^3 \to \mathbb{R}^2$ la función $g(u, v, w) = (u^2 + 2uvw, u + v + w)$. Podemos formar la composición $f \circ g: \mathbb{R}^3 \to \mathbb{R}$, la cual sería $(f \circ g)(u, v, w) = f(g(u, v, w)) = f(u^2+2uvw, u+v+w) = 3(u^2+2uvw)^2+5(u^2+2uvw)(u+v+w)$. Más aún, si consideramos la función $h: \mathbb{R}^2 \to \mathbb{R}^3$, dada por $h(\alpha, \beta) = (\alpha + \beta, \alpha - \beta, \alpha\beta)$, podemos formar la composición $(f \circ g) \circ h: \mathbb{R}^2 \to \mathbb{R}$, que sería $((f \circ g) \circ h)(\alpha, \beta) = (f \circ g)(h(\alpha, \beta)) =$

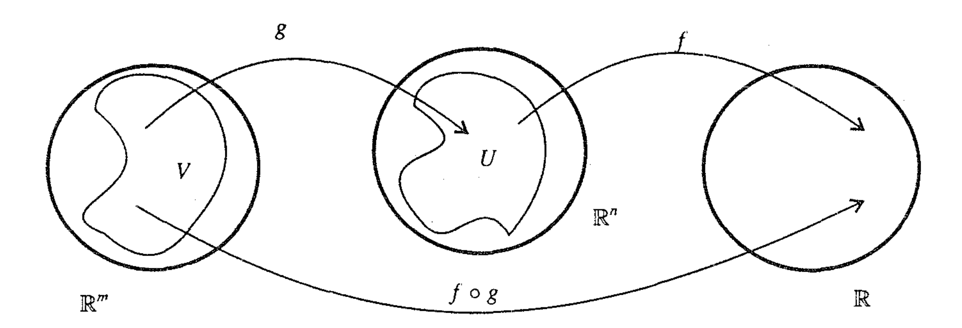


Figura 4. Dominio y codominio de la composición.

$(f \circ g)(\alpha + \beta, \alpha - \beta, \alpha\beta) = 3[(\alpha + \beta)^2 + 2(\alpha + \beta)(\alpha - \beta)(\alpha\beta)]^2 + 5[(\alpha + \beta)^2 + 2(\alpha + \beta)(\alpha - \beta)(\alpha\beta)][\alpha + \beta + \alpha - \beta + \alpha\beta]$. El lector puede verificar que llegaríamos a este mismo resultado haciendo primero la composición $g \circ h: \mathbb{R}^2 \to \mathbb{R}^2$ y luego componiendo con la función f, obteniendo la función $f \circ (g \circ h): \mathbb{R}^2 \to \mathbb{R}$. Es decir que $(f \circ g) \circ h = f \circ (g \circ h)$. Esta función se denota simplemente $f \circ g \circ h$. Esquemáticamente

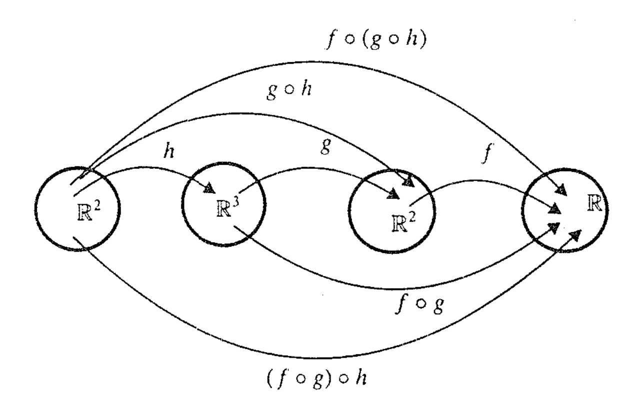


Figura 5. Composición del ejemplo 5.

En general, si tenemos las funciones $f: \mathbb{R}^n \to \mathbb{R}$, $g: \mathbb{R}^m \to \mathbb{R}^n$ y $h: \mathbb{R}^p \to \mathbb{R}^m$, podemos formar las composiciones $f \circ (g \circ h)$, $(f \circ g) \circ h: \mathbb{R}^p \to \mathbb{R}$. Se puede demostrar que estas funciones son iguales, es decir que $(f \circ (g \circ h))(\mathbf{x}) = ((f \circ g) \circ h)(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^p$. Estas funciones se denotan simplemente como $f \circ g \circ h$. Se tiene entonces $(f \circ g \circ h)(\mathbf{x}) = f(g(h(\mathbf{x})))$. ∎

Los conceptos de continuidad y diferenciabilidad para funciones $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $m > 1$, se establecen en términos de las funciones coordenadas de la función f. De modo más preciso, se dirá que la función $f = (f_1, f_2, \ldots, f_m)$ es continua (respectivamente, diferenciable) en el punto $\mathbf{x}_0 \in U$, si y sólo si todas y cada una de las funciones coordenadas $f_i: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, lo son.

Ejemplo 6. La función $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x, |y|)$ es continua en todo su dominio, pues las funciones coordenadas $f_1, f_2: \mathbb{R}^2 \to \mathbb{R}$, $f_1(x, y) = x$, $f_2(x, y) = |y|$ lo son. Esta función no es diferenciable en el origen, pues la función $f_2(x, y) = |y|$ no lo es. Por otra parte, la función $f: \mathbb{R}^3 \to \mathbb{R}^4$, $f(x, y, z) = (x + y + z, \operatorname{sen} x + \cos y, e^{x+z}, \arctan(xz^2))$ es diferenciable en todo su dominio, pues las funciones coordenadas $f_1, f_2, f_3, f_4: \mathbb{R}^3 \to \mathbb{R}$, $f_1(x, y, z) = x + y + z$, $f_2(x, y, z) = \operatorname{sen} x + \cos y$, $f_3(x, y, z) = e^{x+z}$, $f_4(x, y, z) = \arctan(xz^2)$ lo son. ∎

Ejercicios (Capítulo 3, Sección 1)

En cada uno de los ejercicios 1–5 se da una función $f: \mathbb{R}^2 \to \mathbb{R}$ y dos funciones $g_1, g_2: \mathbb{R}^2 \to \mathbb{R}$. **a.** Escriba la función $g: \mathbb{R}^2 \to \mathbb{R}^2$ que tenga por funciones coordenadas a g_1 y g_2. **b.** Determine la función compuesta $F(u, v) = (f \circ g)(u, v)$

1. $f(x, y) = x$, $g_1(u, v) = u$, $g_2(u, v) = v$

2. $f(x, y) = \operatorname{sen} x + \operatorname{sen} y$, $g_1(u, v) = u^2 v$, $g_2(u, v) = uv$
3. $f(x, y) = 3x^2 + 8y^3$, $g_1(u, v) = u + v$, $g_2(u, v) = u$
4. $f(x, y) = \dfrac{x + y}{x^2 + y^2 + 1}$, $g_1(u, v) = \operatorname{sen} u$, $g_2(u, v) = \cos u$
5. $f(x, y) = \arctan^{35}(x + y)$, $g_1(u, v) = u - v + \pi/4$, $g_2(u, v) = v - u$

En los ejercicios 6–10 se da una función $f: \mathbb{R}^2 \to \mathbb{R}$. Determine las funciones: **a.** $F(x, y) = f(f(x, y), y)$; **b.** $F(x, y) = f(x, f(x, y))$; **c.** $F(x, y) = f(f(x, y), f(x, y))$; **d.** $F(x, y) = f(f(x, y), f(y, x))$

6. $f(x, y) = ax + by$
7. $f(x, y) = x^2 y$
8. $f(x, y) = \operatorname{sen} x + \operatorname{sen} y$
9. $f(x, y) = \ln(1 + |x|)$
10. $f(x, y) = \arctan y$
11. Sea $f: \mathbb{R}^n \to \mathbb{R}$ la función lineal $f(x_1, x_2, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + b$. Demuestre que si componemos esta función con n funciones lineales $g_i: \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, n$, la función compuesta $F: \mathbb{R}^n \to \mathbb{R}$, $F(x) = f(g_1(x), g_2(x), \ldots, g_n(x))$ también es lineal. Compruebe que el término independiente de F es $F(0) = \sum_{i=1}^{n} a_i g_i(0)$.
12. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ la función $f(x_1, x_2, \ldots, x_n) = \frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}$. Determine la función compuesta $F(x_1, x_2, \ldots, x_n) = f(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n})$. ¿Dónde está definida esta función? (La respuesta NO es "en todo $\mathbb{R}^n$").

En los ejercicios 13–20 se dan funciones f, g, h definidas en algún conjunto de $\mathbb{R}^n$ (algún n) y tomando valores en $\mathbb{R}^m$ (algún m). Identifique en cada caso los espacios $\mathbb{R}^n$ y $\mathbb{R}^m$ del dominio y codominio de cada función, y, *cuando sea posible*, determine las funciones compuestas: **a.** $f \circ g$; **b.** $h \circ g$; **c.** $f \circ g \circ h$.

13. $f(x, y) = x + 2y$, $g(u, v) = (3u, v)$, $h(r, s) = r + 4s$
14. $f(x, y) = xy$, $g(u, v) = (uv, u^2 v^2)$, $h(r, s) = r$
15. $f(x, y) = x^2 + y^2$, $g(u, v) = (u + v, v - u)$, $h(t) = (t, 2t)$
16. $f(x, y) = 5x + 3y$, $g(u) = \operatorname{sen} u$, $h(r, s) = \left(\dfrac{r}{s}, \dfrac{s}{r}\right)$
17. $f(x, y, z) = x + z$, $g(u) = (u, 2u, 3u)$, $h(t) = t^2$
18. $f(x, y, z) = xyz$, $g(u, v, w) = (uv, uw, vw)$, $h(r, s, t) = (r^2, s^3, t^4)$
19. $f(x, y, z) = \operatorname{sen} x$, $g(u, v, w) = (u, v^{34}, uv \arctan^5 w)$, $h(t) = (t, t^3, t + 4)$
20. $f(x, y, z) = 1$, $g(u, v, w) = (w, v, u)$, $h(r, s, t) = (r^2, rst, s^5 t^4)$
21. Determine funciones $f: \mathbb{R}^2 \to \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$ de modo que la función dada $F: \mathbb{R}^2 \to \mathbb{R}$ se vea como $F = g \circ f$
    a. $F(x, y) = \operatorname{sen}(3xy)$
    b. $F(x, y) = \arctan^2(x^3 - \operatorname{sen} y)$
    c. $F(x, y) = \sqrt{1 - x^2 - y^2}$
    d. $F(x, y) = \ln(x - y)$

22. Determine funciones $f: \mathbb{R}^2 \to \mathbb{R}$, $g: \mathbb{R}^2 \to \mathbb{R}^2$ de modo que la función dada $F: \mathbb{R}^2 \to \mathbb{R}$ se vea como $F = f \circ g$

a. $F(x, y) = \operatorname{sen}(x + y) + \cos(x - y)$

b. $F(x, y) = \arctan^2(5x + y^2) + 3x - 2y$

c. $F(x, y) = (3x - y)^3 + 4(2x + y)^4$

d. $F(x, y) = \ln|xy| + e^x$

23. Determine funciones $f: \mathbb{R}^3 \to \mathbb{R}$, $g: \mathbb{R}^3 \to \mathbb{R}^3$ de modo que la función dada $F: \mathbb{R}^3 \to \mathbb{R}$ se vea como $F = f \circ g$

a. $F(x, y, z) = x^2 + y^2 + z^2$

b. $F(x, y, z) = \operatorname{sen}(x + y) + \cos(y - z) + \operatorname{sen}^2(x + y + z)$

c. $F(x, y, z) = (x + y + z)^2 + 5(3x - y)^4 - 10z$

d. $F(x, y, z) = \ln(1 + x^2 + y^2 + z^2) + x + y + z$

24. Determine funciones $f: \mathbb{R}^3 \to \mathbb{R}$, $g_i: \mathbb{R} \to \mathbb{R}$, $i = 1, 2, \ldots, 13$, de modo que $F = g_{13} \circ g_{12} \circ \cdots \circ g_2 \circ g_1 \circ f$, si

$$F(x, y, z) = 1 + 5\sqrt{5 + \ln^2(1 + 8 \operatorname{sen}^4 \sqrt{3 + x^2y^2z^2})}$$

25. Considere las funciones $f, g: \mathbb{R}^2 \to \mathbb{R}^2$ dadas por

$$f(x, y) = (x + y, 2x + 3y), \qquad g(x, y) = (3x - y, -2x + y)$$

a. Determine $f(2, 4)$ y $g(f(2, 4))$

b. Determine $g(2, 4)$ y $f(g(2, 4))$

c. Demuestre que las composiciones $f \circ g$ y $g \circ f$ son iguales a la función identidad $\mathrm{id}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathrm{id}(x, y) = (x, y)$.

26. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (y, x)$. ¿A qué es igual la función compuesta $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F = f \circ f$?

27. Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (2x - 3y - 5z, -x + 4y + 5z, x - 3y - 4z)$. Determine la función $F: \mathbb{R}^3 \to \mathbb{R}^3$, $F = f \circ f$.

28. Dé ejemplos de funciones $f, g: \mathbb{R}^2 \to \mathbb{R}^2$ tales que: **a.** $f \circ g = g \circ f$; **b.** $f \circ g \neq g \circ f$.

29. Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (x + y + 3z, 5x + 2y + 6z, -2x - y - 3z)$. Determine la función $F: \mathbb{R}^3 \to \mathbb{R}^3$, $F = f \circ f \circ f$.

30. Se dice que la función $f: \mathbb{R}^2 \to \mathbb{R}^2$ es lineal (en $\mathbb{R}^2$), si sus funciones coordenadas $f_1, f_2: \mathbb{R}^2 \to \mathbb{R}$ lo son. Es decir, f es lineal si es de la forma $f(x, y) = (a_1x + b_1y + c_1, a_2x + b_2y + c_2)$. Sea f una función lineal en $\mathbb{R}^2$ tal que $f(0, 0) = (0, 0)$. Demuestre que $f(\mathbf{p} + \alpha\mathbf{q}) = f(\mathbf{p}) + \alpha f(\mathbf{q})$, donde $\mathbf{p}, \mathbf{q} \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$. Compruebe que las funciones de los ejercicios 25 y 26 son de este tipo.

31. Sea $g: \mathbb{R} \to \mathbb{R}^n$ una función continua en $x_0 \in \mathbb{R}$, y $f: \mathbb{R}^n \to \mathbb{R}$ una función continua en $g(x_0) \in \mathbb{R}^n$. Demuestre que la función $f \circ g: \mathbb{R} \to \mathbb{R}$ es continua en x_0.

(*)32. Sea $f: \mathbb{R}^n \to \mathbb{R}^m$ una función definida en $\mathbb{R}^n$. Sea $\mathbf{x}_0 \in \mathbb{R}^n$. Demuestre que las siguientes afirmaciones acerca de la función f son equivalentes:

a. f es continua en $\mathbf{x}_0$.

b. dado un $\epsilon > 0$ existe un $\delta > 0$ tal que

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \to \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \epsilon$$

c. la imagen inversa de cualquier conjunto abierto (en $\mathbb{R}^m$) que contenga a $f(\mathbf{x}_0)$ es un conjunto abierto (en $\mathbb{R}^n$) que contiene a $\mathbf{x}_0$. Es decir, si llamamos $B_{f(\mathbf{x}_0)}$ a un conjunto abierto que contiene a $f(\mathbf{x}_0)$, entonces (demuestre que) f es continua en $\mathbf{x}_0$ si y solamente si

$$f^{-1}(B_{f(\mathbf{x}_0)}) = \{\mathbf{x} \in \mathbb{R}^n | f(\mathbf{x}) \in B_{f(\mathbf{x}_0)}\} \subseteq \mathbb{R}^n$$

es un conjunto abierto. (Ver ejercicio 83 de la sección 3 del capítulo 2).

3.2 Regla de la cadena

Ahora estudiaremos la relación entre la diferenciabilidad y las derivadas parciales de una función compuesta, con la diferenciabilidad y las derivadas parciales de sus funciones componentes.

Enunciaremos rigurosamente el teorema que establece la regla de la cadena para la derivación y presentamos (de manera opcional) una demostración en un caso concreto. Quisiéramos advertir, sin embargo, que en este tema, como sucede en algunos otros tópicos del cálculo, existen dos aspectos que demandan atención y esfuerzo por parte del lector para su comprensión: uno de ellos es el aspecto "teórico" que está contenido en el teorema que presentamos y , desde luego, su demostración. El otro aspecto es el carácter "práctico": en él no interesará mucho escribir en detalle las fórmulas de la regla de la cadena que aparecerán en el teorema, sino más bien adquirir habilidad en su uso para aplicarlas adecuadamente en distintas situaciones. Creemos que, en un primer acercamiento a este tema, el segundo aspecto mencionado es sobre el que se debe poner atención especial y a él se dedicará una buena cantidad de ejemplos en esta sección.

Teorema 3.2.1 (Regla de la cadena). Sea $g: V \subseteq \mathbb{R}^m \to \mathbb{R}^n$ una función definida en el conjunto abierto V de $\mathbb{R}^m$, diferenciable en $\mathbf{x}_0 \in V$ (esto significa que las n funciones coordenadas de g, digamos $g_i: V \subseteq \mathbb{R}^m \to \mathbb{R}$, $i = 1, 2, \ldots, n$ son diferenciables en $\mathbf{x}_0 \in V$). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$, tal que $g(V) \subseteq U$, diferenciable en el punto $g(\mathbf{x}_0) \in U$. Entonces la composición $f \circ g: V \subseteq \mathbb{R}^m \to \mathbb{R}$ es diferenciable en $\mathbf{x}_0$ y sus derivadas parciales son

$$\frac{\partial}{\partial x_j}(f \circ g)(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(g(\mathbf{x}_0))\frac{\partial g_i}{\partial x_j}(\mathbf{x}_0), \qquad j = 1, 2, \ldots, m$$

Demostración. (OPCIONAL, caso $m = n = 2$). Se tiene la situación siguiente (figura 1) en que las dos funciones coordenadas de g, $g_{1\,2}: V \subseteq \mathbb{R}^2 \to \mathbb{R}$ son diferenciables en $\mathbf{x}_0$ y f es diferenciable en $g(\mathbf{x}_0)$. Los puntos del dominio de g se denotan con coordendas x_i y los del dominio de f con coordenadas y_i. Así, $\mathbf{x}_0 = (x_1, x_2) \in V$ y $g(\mathbf{x}_0) = (y_1, y_2)$.

La noción de diferenciabilidad que manejaremos en la hipótesis de esta demostración es como sigue: la función $\varphi: B \subseteq \mathbb{R}^2 \to \mathbb{R}$ se dirá diferenciable en el punto $\mathbf{x}_0$ del conjunto abierto B de $\mathbb{R}^2$, si al escribir

$$\varphi(\mathbf{x}_0 + \mathbf{H}) = \varphi(\mathbf{x}_0) + \frac{\partial \varphi}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial \varphi}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho(\mathbf{H})$$

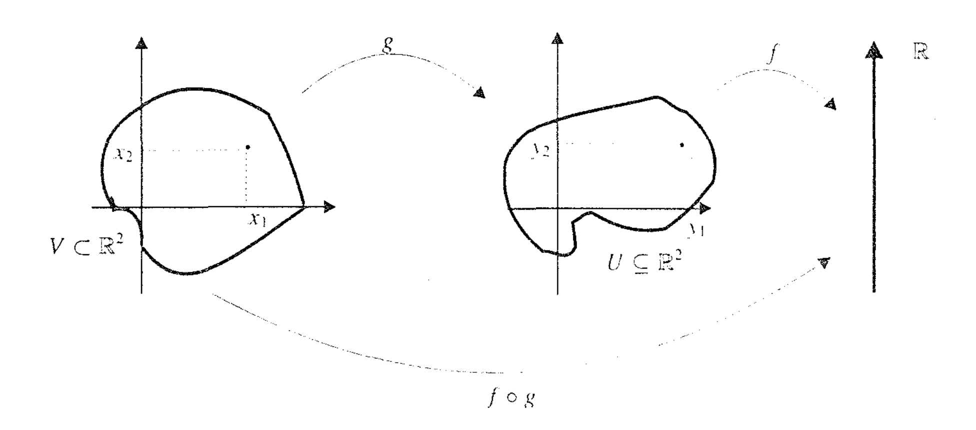


Figura 1. Las funciones f y g

donde $\mathbf{H} = (h_1, h_2)$ es $\mathbf{x}_0 + \mathbf{H} \in B$, la función ρ definida en alguna bola abierta de $\mathbb{R}^2$ con centro en $\mathbf{0}$, es tal que $\rho(\mathbf{0}) = 0$ y es continua en $\mathbf{0}$, es decir, $\lim_{\mathbf{H}\to\mathbf{0}} \rho(\mathbf{H}) = 0$. El lector puede verificar la equivalencia de esta definición con la dada en el capítulo anterior (ver ejercicio 29 de la sección 6, capítulo 2).

Nuestra hipótesis en el teorema será entonces:

i. La función $g_1: V \subseteq \mathbb{R}^2 \to \mathbb{R}$ (la primera función coordenada de g) es diferenciable en $\mathbf{x}_0 \in V$. Es decir

$$g_1(\mathbf{x}_0 + \mathbf{H}) = g_1(\mathbf{x}_0) + \frac{\partial g_1}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_1(\mathbf{H})$$

donde $\rho_1(\mathbf{0}) = 0$ y $\lim_{\mathbf{H}\to\mathbf{0}} \rho_1(\mathbf{H}) = 0$.

ii. La función $g_2: V \subseteq \mathbb{R}^2 \to \mathbb{R}$ (la segunda función coordenada de g) es diferenciable en $\mathbf{x}_0 \in V$. Es decir

$$g_2(\mathbf{x}_0 + \mathbf{H}) = g_2(\mathbf{x}_0) + \frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_2(\mathbf{H})$$

donde $\rho_2(\mathbf{0}) = 0$ y $\lim_{\mathbf{H}\to\mathbf{0}} \rho_2(\mathbf{H}) = 0$.

iii. La función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en $g(\mathbf{x}_0)$. Es decir

$$f(g(\mathbf{x}_0) + \mathbf{K}) = f(g(\mathbf{x}_0)) + \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))k_1 + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))k_2 + \|\mathbf{K}\|\rho_3(\mathbf{K})$$

donde $\rho_3(\mathbf{0}) = 0$ y $\lim_{K\to 0} \rho_3(\mathbf{K}) = 0$.

Nuestro trabajo consistirá en probar la diferenciabilidad de la función $f \circ g$ en $\mathbf{x}_0$ (en el sentido de la definición del capítulo anterior). Es decir, que si escribimos

$$(f \circ g)(\mathbf{x}_0 + \mathbf{H}) = (f \circ g)(\mathbf{x}_0) + Ah_1 + Bh_2 + r(\mathbf{H}) \ldots \qquad *$$

(donde en A y B deberán aparecer las derivadas parciales de $f \circ g$ respecto de x_1 y x_2, respectivamente), se debe tener que

$$\lim_{\mathbf{H}\to 0} \frac{r(\mathbf{H})}{\|\mathbf{H}\|} = 0$$

Comencemos:

$$(f \circ g)(\mathbf{x}_0 + \mathbf{H}) = f(g(\mathbf{x}_0 + \mathbf{H})) = f(g_1(\mathbf{x}_0 + \mathbf{H}), g_2(\mathbf{x}_0 + \mathbf{H}))$$

usando **i** y **ii** nos queda

$$\begin{aligned}(f \circ g)(\mathbf{x}_0 + \mathbf{H}) &= f\Bigg(g_1(\mathbf{x}_0) + \frac{\partial g_1}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_1(\mathbf{H}),\\ &\qquad g_2(\mathbf{x}_0) + \frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_2(\mathbf{H})\Bigg)\\ &= f\Bigg(\big(g_1(\mathbf{x}_0), g_2(\mathbf{x}_0)\big) + \Bigg(\frac{\partial g_1}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_1(\mathbf{H}),\\ &\qquad \frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_2(\mathbf{H})\Bigg)\Bigg)\end{aligned}$$

Llamemos

$$\mathbf{K} = (k_1, k_2) = \left(\frac{\partial g_1}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_1(\mathbf{H}), \frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_2(\mathbf{H})\right)$$

Entonces

$$(f \circ g)(\mathbf{x}_0 + \mathbf{H}) = f(g(\mathbf{x}_0) + \mathbf{K})$$

Al utilizar **iii**, nos queda

$$(f \circ g)(\mathbf{x}_0 + \mathbf{H}) = f(g(\mathbf{x}_0)) + \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))k_1 + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))k_2 + \|\mathbf{K}\|\rho_3(\mathbf{K})$$

Haciendo explícitas las coordenadas de $\mathbf{K}$, k_1 y k_2, nos queda

$$\begin{aligned}(f \circ g)(\mathbf{x}_0 + \mathbf{H}) &= f(g(\mathbf{x}_0)) + \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\left(\frac{\partial g_1}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_1(\mathbf{H})\right)\\ &\quad + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\left(\frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)h_2 + \|\mathbf{H}\|\rho_2(\mathbf{H})\right) + \|\mathbf{K}\|\rho_3(\mathbf{K})\\ &= f(g(\mathbf{x}_0)) + \left(\frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\frac{\partial g_1}{\partial x_1}(\mathbf{x}_0) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)\right)h_1\\ &\quad + \left(\frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\frac{\partial g_1}{\partial x_2}(\mathbf{x}_0) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)\right)h_2\\ &\quad + \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\|\mathbf{H}\|\rho_1(\mathbf{H}) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\|\mathbf{H}\|\rho_2(\mathbf{H}) + \|\mathbf{K}\|\rho_3(\mathbf{K})\end{aligned}$$

Comparando esta última expresión con (*) podemos escribir que

$$\begin{aligned}A &= \frac{\partial}{\partial x_1}(f \circ g)(\mathbf{x}_0) = \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\frac{\partial g_1}{\partial x_1}(\mathbf{x}_0) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\frac{\partial g_2}{\partial x_1}(\mathbf{x}_0)\\ B &= \frac{\partial}{\partial x_2}(f \circ g)(\mathbf{x}_0) = \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\frac{\partial g_1}{\partial x_2}(\mathbf{x}_0) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\frac{\partial g_2}{\partial x_2}(\mathbf{x}_0)\\ r(\mathbf{H}) &= \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\|\mathbf{H}\|\rho_1(\mathbf{H}) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\|\mathbf{H}\|\rho_2(\mathbf{H}) + \|\mathbf{K}\|\rho_3(\mathbf{K})\end{aligned}$$

en lo que se establecen las fórmulas de las derivadas parciales de la composición $f \circ g$ respecto de x_1 y x_2. Queda por demostrar que el residuo $r(\mathbf{H})$ cumple con la propiedad requerida. Se tiene

$$\frac{r(\mathbf{H})}{\|\mathbf{H}\|} = \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\rho_1(\mathbf{H}) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\rho_2(\mathbf{H}) + \frac{\|\mathbf{K}\|}{\|\mathbf{H}\|}\rho_3(\mathbf{K})$$

Es un pequeño trabajo técnico, que dejamos al lector, probar que la cantidad $\frac{\|\mathbf{K}\|}{\|\mathbf{H}\|}$ se mantiene limitada para $\mathbf{H}$ en una bola con centro en el origen de $\mathbb{R}^2$. Siendo así, al tomar límite cuando $\mathbf{H} \to \mathbf{0}$, vemos que los dos primeros sumandos tienden a cero pues, ρ_1 y ρ_2 tienen dicha propiedad, y el tercer sumando, al mantenerse limitado el factor $\frac{\|\mathbf{K}\|}{\|\mathbf{H}\|}$ y tender ρ_3 a cero (pues si $\mathbf{H} \to \mathbf{0}$ también $\mathbf{K} \to \mathbf{0}$, y $\rho_3(\mathbf{K})$ tiene la propiedad de $\lim_{\mathbf{K}\to 0} \rho_3(\mathbf{K}) = 0$) también tiende a cero. Con esto probamos que $\lim_{\mathbf{H}\to 0} \frac{r(\mathbf{H})}{\|\mathbf{H}\|} = 0$ como queríamos. Q.E.D.

La fórmula del teorema anterior adquiere una apariencia especial si la función g depende de una sola variable. Más en concreto, si tenemos la situación siguiente

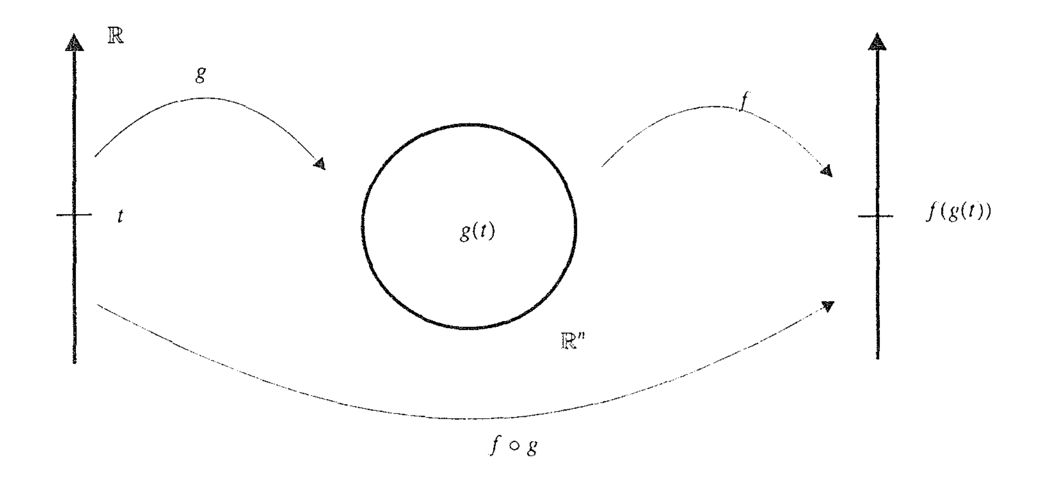


Figura 2. Caso especial de la regla de la cadena.

Vemos que g es una función de n funciones coordenadas, cada una de ellas dependiendo de una sola variable, digamos t. Al hacer la composición con f, se obtiene también la función real $f \circ g$ que depende sólo de t. Para estas funciones, las derivadas parciales de las fórmulas del teorema anterior son *derivadas totales*, quedándonos como

$$\frac{d}{dt}(f \circ g)(t) = (f \circ g)'(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(g(t))\frac{dg_i}{dt}(t)$$

donde g_i, $i = 1, 2, \ldots, n$, son las funciones coordendas de g.

Antes de empezar a ver algunos ejemplos, veamos la demostración ya prometida desde la sección 8 del capítulo anterior del hecho de que el gradiente de una función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto $\mathbf{x}_0 \in U$, es ortogonal a la curva (o superficie) de nivel de f que pasa por $\mathbf{x}_0$, hecho que

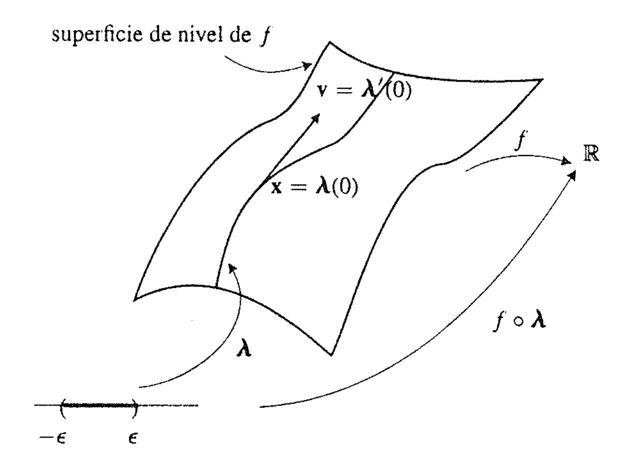


Figura 3. Grafica del nivel c de f.

ya hemos usado en repetidas ocasiones y del que se había dado un argumento geométrico que lo validaba.

Consideremos pues la función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$ y sea $\mathbf{x}_0 \in U$. Sea $c = f(\mathbf{x}_0)$. La curva (o superficie) de nivel que pasa por $\mathbf{x}_0$ (el nivel c de f) es

$$S = \{\mathbf{x} \in \mathbb{R}^n | f(\mathbf{x}) = c\}$$

Sea $\bar{\mathbf{x}}$ un punto cualquiera de S y sea $\mathbf{v} \in \mathbb{R}^n$ un vector tangencial cualquiera a S en $\bar{\mathbf{x}}$. Queremos ver que el vector grad $f(\bar{\mathbf{x}})$ es ortogonal a $\mathbf{v}$. (Como $\bar{\mathbf{x}}$ y $\mathbf{v}$ son arbitrarios, esto probaría lo que deseamos: que el vector grad $f(\mathbf{x})$ es ortogonal a la curva de superficie de nivel que pasa por $\mathbf{x}$.) Tomemos una función $\lambda: I = (-\epsilon, \epsilon) \subseteq \mathbb{R} \to \mathbb{R}^n$ tal que $\lambda(0) = \bar{\mathbf{x}}$, $\lambda'(0) = \mathbf{v}$, y $\lambda(t) \in S, \forall t \in I$. Una función de este tipo se llama "curva en $\mathbb{R}^n$". El estudio de estas funciones se efectuará en el capítulo 5. Nuestra curva λ que estamos considerando es tal que su imagen cae completamente sobre el nivel c de f, pasando en $t = 0$ por el punto $\bar{\mathbf{x}}$ y teniendo ahí por derivada (la cual, como veremos luego, es un vector tangente a la curva) al vector $\mathbf{v}$. Consideremos que λ es diferenciable, lo que significa que si $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, las funciones coordenadas $\lambda_i: (-\epsilon, \epsilon) \to \mathbb{R}$ son funciones diferenciables. Esquemáticamente

Podemos formar la composición $f \circ \lambda: (-\epsilon, \epsilon) \to \mathbb{R}$. Se observa que $(f \circ \lambda)(t) = f(\lambda(t)) = c$, por estar $\lambda(t)$ siempre en S. Según la regla de la cadena tenemos, derivando, que

$$\frac{\partial}{\partial t}(f \circ \lambda)(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{\mathbf{x}}) \frac{\partial \lambda_i}{\partial t}(0) = 0$$

Podemos ver el lado izquierdo de esta expresión como el producto punto del vector grad $f(\bar{\mathbf{x}}) = \frac{\partial f}{\partial x_1}(\bar{\mathbf{x}}), \ldots, \frac{\partial f}{\partial x_n}(\bar{\mathbf{x}})$ y el vector $\boldsymbol{\lambda}'(0) = (\lambda_1'(0), \ldots, \lambda_n'(0)) = \mathbf{v}$. Hemos entonces probado que

$$\text{grad } f(\bar{\mathbf{x}}) \cdot \mathbf{v} = 0$$

lo que muestra la ortogonalidad de grad $f(\bar{\mathbf{x}})$ con $\mathbf{v}$, como queríamos.

Veamos ahora algunos ejemplos.

Ejemplo 1. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ la función dada por $f(x, y) = x^2 + 3y^2$, y sea $g: U \subseteq \mathbb{R} \to \mathbb{R}^2$ la función $g(t) = (e^t, \cos t)$. Podemos, sin hacer explícita la composición $f \circ g: \mathbb{R} \to \mathbb{R}$, calcular sus derivadas parciales usando la fórmula dada en el teorema anterior

$$\begin{aligned}\frac{\partial}{\partial t}(f \circ g)(t) &= \frac{\partial f}{\partial x}(g(t))\frac{\partial}{\partial t}(e^t) + \frac{\partial f}{\partial y}(g(t))\frac{\partial}{\partial t}(\cos t)\\ &= 2x\big|_{g(t)}(e^t) + 6y\big|_{g(t)}(-\operatorname{sen} t) =\\ &= 2e^te^t + 6\cos t(-\operatorname{sen} t) = 2e^{2t} - 6\cos t \operatorname{sen} t\end{aligned}$$

Se puede llegar a este mismo resultado haciendo primero la composición $(f \circ g)(t) = f(g(t)) = f(e^t, \cos t) = e^{2t} + 3\cos^2 t$, y derivando directamente en ella, como se comprueba fácilmente. ■

Ejemplo 2. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función tal que $\operatorname{grad} f(1, 1) = (2, 4)$ y $g: \mathbb{R}^3 \to \mathbb{R}^2$ una función tal que sus funciones coordenadas $g_i: \mathbb{R}^3 \to \mathbb{R}$, $i = 1, 2$, tienen los siguientes gradientes: $\operatorname{grad} g_1(1, 1, 1) = (2, -3)$, $\operatorname{grad} g_2(1, 1, 1) = (-5, 4)$. Si $g(1, 1, 1) = (1, 1)$, se quiere hallar $\frac{\partial}{\partial x}(f \circ g)(1, 1, 1)$. Según la regla de la cadena (llamando (x, y, z) a los puntos de $\mathbb{R}^3$ y (u, v) a los de $\mathbb{R}^2$) se tiene.

$$\begin{aligned}\frac{\partial}{\partial x}(f \circ g)(1, 1, 1) &= \frac{\partial f}{\partial u}(g(1, 1, 1))\frac{\partial g_1}{\partial x}(1, 1, 1) + \frac{\partial f}{\partial v}(g(1, 1, 1))\frac{\partial g_2}{\partial x}(1, 1, 1)\\ &= \frac{\partial f}{\partial u}(1, 1)\frac{\partial g_1}{\partial x}(1, 1, 1) + \frac{\partial f}{\partial v}(1, 1)\frac{\partial g_2}{\partial x}(1, 1, 1)\\ &= (2)(2) + (4)(-5) = -16.\end{aligned}$$

■

Una vez establecido el teorema importante de esta sección, con la notación y rigor adecuados, pasemos a la parte "operativa": ahora nos interesa sacar provecho del teorema para hacer cálculo de derivadas de ciertas funciones compuestas. La situación típica que se presenta en la práctica (donde se suele ser poco riguroso, incluso en la notación), es la siguiente: si tenemos una función $z = z(y_1, y_2, \ldots, y_n)$ de n variables y cada una de ellas es a su vez función de otras m variables, digamos $y_i = y_i(x_1, x_2, \ldots, x_m)$ (observe que —como es normal en estos teoremas— usamos las mismas letras "z", "y" para denotar a las funciones $z: \mathbb{R}^n \to \mathbb{R}$, $y: \mathbb{R}^m \to \mathbb{R}$ así como a sus imágenes), entonces podemos formar la composición

$$z\big(y_1(x_1, \ldots, x_m), y_2(x_1, \ldots, x_m), \ldots, y_n(x_1, \ldots, x_m)\big)$$

de la cual nos interesa obtener las derivadas parciales respecto de las variables x_i.

Atención: suele ser muy común denotar esta función como

$$z = z(y_1(x_1, \ldots, x_m), y_2(x_1, \ldots, x_m), \ldots, y_n(x_1, \ldots, x_m))$$

y escribir $\frac{\partial z}{\partial x_i}$, $i = 1, \ldots, m$, para la derivada parcial de ella respecto de x_i. Esto puede ser motivo de confusiones (ver ejercicio 22 al final de la sección). Sin embargo, tiene la "ventaja" de hacer aún más simple la notación en el proceso del cálculo de las derivadas parciales de la función compuesta

funciones dadas

$$z = z(y_1, y_2, \ldots, y_n) \qquad \begin{aligned} y_1 &= y_1(x_1, x_2, \ldots, x_m) \\ y_2 &= y_2(x_1, x_2, \ldots, x_m) \\ &\vdots \\ y_n &= y_n(x_1, x_2, \ldots, x_m) \end{aligned}$$

↓ se componen

$$z = z\big(y_1(x_1, x_2, \ldots, x_m), y_2(x_1, x_2, \ldots, x_m), \ldots, y_n(x_1, x_2, \ldots, x_m)\big)$$

función compuesta

Según la regla de la cadena se tiene entonces

$$\boxed{\frac{\partial z}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial z}{\partial y_i}\frac{\partial y_i}{\partial x_j}} \qquad j = 1, 2, \ldots, m$$

No podemos negar la ventajosa simplicidad de la fórmula de las derivadas parciales de la función compuesta con esta presentación, respecto de su análoga que apareció en el teorema anterior. Su ventaja es clara: nos permitirá una mecanización operativa fácil en el cálculo de derivadas y nos permitirá "ver" claramente el proceso involucrado en el cálculo Sin embargo, debemos hacer notar que: 1. no se hacen explícitos los puntos donde están calculadas las derivadas; 2. se usan letras iguales para denotar cosas distintas, a saber

$$\frac{\partial \overset{\overset{1}{\downarrow}}{z}}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial \overset{\overset{2}{\downarrow}}{z}}{\partial \underset{\underset{4}{\uparrow}}{y_i}} \cdot \frac{\partial \overset{\overset{3}{\downarrow}}{y_i}}{\partial x_j}$$

1 se refiere a la función compuesta de variables $x_1, x_2, \ldots, x_n$
2 se refiere a la función dada
3 se refiere a la i-ésima función dada con que se compone la función z
4 se refiere a la i-ésima variable de la función dada

Estos hechos se deben poder manejar naturalmente después de una buena cantidad de ejercicios.

Ejemplo 3. Si $z = x^2 + 3y^2$, donde $x = e^t$, $y = \cos t$ (ver ejemplo 1), entonces

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} = 2xe^t - 6y \operatorname{sen} t = 2e^te^t - 6\cos t \operatorname{sen} t =$$
$$= 2e^{2t} - 6\cos t \operatorname{sen} t$$ ■

Ejemplo 4. Si $z = x^2 e^{y^3}$, donde $x = uv$, $y = u^2 - v^3$ entonces

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} = 2xe^{y^3}v + 3x^2y^2e^{y^3}(2u)$$
$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v} = 2xe^{y^3}u + 3x^2y^2e^{y^3}(-3v^2)$$ ■

Ejemplo 5. Si $u = 3x^2y^2z^2 + 6\,\text{sen}(xyz)$, donde

$$x = 4vw^2, \quad y = 5v^2 + 10w^3, \quad z = v^3$$

Entonces

$$\begin{aligned}\frac{\partial u}{\partial v} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial v} = \\ &= (6xy^2z^2 + 6yz\cos(xyz))(4w^2) + (6x^2yz^2 + 6xz\cos(xyz))(10v) \\ &\quad + (6x^2y^2z + 6xy\cos(xyz))(3v^2) \\ \frac{\partial u}{\partial w} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial w} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial w} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial w} \\ &= (6xy^2z^2 + 6yz\cos(xyz))(8vw) + (6x^2yz^2 + 6xz\cos(xyz))(30w^2) \\ &\quad + (6x^2y^2z + 6xy\cos(xyz))(0)\end{aligned}$$

■

Ejemplo 6. Consideremos la función $z = f(x^2 + y^3, 2x^2 - y^2)$. Obsérvese que ésta se presenta ya como la composición de $z = f(u, v)$ con las funciones $u = x^2 + y^3$, $v = 2x^2 - y^2$. Se tiene entonces

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = 2x\frac{\partial f}{\partial u} + 4x\frac{\partial f}{\partial v} \\ \frac{\partial z}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} = 3y^2\frac{\partial f}{\partial u} - 2y\frac{\partial f}{\partial v}\end{aligned}$$

Llamamos la atención al hecho de que las letras x, y, z, u, v, w no tienen (en estos temas) "posiciones privilegiadas" para representar las variables de algunas de las funciones que componen (o bien de la función compuesta). En este ejemplo, u, v juegan los papeles contrarios a x, y respecto del ejemplo 4. ■

Ejemplo 7. La función $z = f\left(\frac{x}{y}\right)$ se debe interpretar como la composición de la función de una sola variable $z = z(u)$ con la función de dos variables $u = \frac{x}{y}$ (quedando así la composición como función de las variables x y y). Entonces

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} = f'(u)\left(\frac{1}{y}\right) \\ \frac{\partial z}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} = f'(u)\left(-\frac{x}{y^2}\right)\end{aligned}$$

■

Ejemplo 8. Decir que una función *satisface* una ecuación significa que es *solución* de la ecuación, entendiéndose "solución" en el sentido clásico: al sustituir "la solución" en la ecuación, ésta queda reducida a una identidad. Por ejemplo en la ecuación algebraica $x^3 + 3x^2 - 4x^2 = 0$, tenemos que $x = 1$ es una solución, pues $(1)^3 + 3(1)^2 - 4(1)^2 = 0$ es una identidad (otra solución es $x = 0$). Del mismo modo, en la ecuación diferencial ordinaria $y'' - y' - 2y = 0$, la función $y = e^{2t}$ es una solución, pues al sustituir esta función en la ecuación obtenemos $(e^{2t})'' - (e^{2t})' - 2e^{2t} = 4e^{2t} - 2e^{2t} - 2e^{2t} = 0$ que es una identidad (otra solución es $y = e^t$, verifique). De la misma manera, podemos comprobar

que la función $z = f(x+y, -x-y)$ donde f es una función diferenciable cualquiera (de $\mathbb{R}^2$ en $\mathbb{R}$) es solución de la ecuación diferencial parcial

$$\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y} = 0$$

En efecto, poniendo $u = x + y$, $v = -x - y$ se tiene

$$\frac{\partial z}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial f}{\partial u}(1) + \frac{\partial f}{\partial v}(-1) = \frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}$$
$$\frac{\partial z}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} = \frac{\partial f}{\partial u}(1) + \frac{\partial f}{\partial v}(-1) = \frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}$$

de modo que al sustituir estas derivadas en la ecuación se obtiene

$$\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y} = \frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} - \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v} = 0$$

que es una identidad. ■

Ejemplo 9. (De nuevo el Teorema de Euler sobre funciones homogéneas). En el apéndice de la sección 7 del capítulo anterior (en el corolario del teorema ahí presentado) vimos que si $f: \mathbb{R}^n \to \mathbb{R}$ es una función homogénea de grado α, digamos $z = f(x_1, x_2, \ldots, x_n)$, entonces ésta satisface la relación

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n} = \alpha f$$

Veremos ahora otro argumento que muestra este hecho, usando las ideas estudiadas en esta sección. Siendo f una función homogénea de grado α tenemos que

$$f(tx) = f(tx_1, tx_2, \ldots, tx_n) = t^\alpha f(x_1, x_2, \ldots, x_n), \qquad t > 0$$

La expresión $f(tx_1, tx_2, \ldots, tx_n)$ se puede ver como la composición de la función $f(u_1, u_2, \ldots, u_n)$ con las n funciones $u_i = tx_i$, $i = 1, 2, \ldots, n$, que son funciones de las variables t y x_i. Derivando entonces la igualdad $f(tx) = f(tx_1, tx_2, \ldots, tx_n) = t^\alpha f(x_1, x_2, \ldots, x_n)$ *respecto de* t, obtenemos

$$\frac{\partial f}{\partial u_1}\frac{\partial u_1}{\partial t} + \frac{\partial f}{\partial u_2}\frac{\partial u_2}{\partial t} + \cdots + \frac{\partial f}{\partial u_n}\frac{\partial u_n}{\partial t} = \alpha t^{\alpha-1} f(x_1, x_2, \ldots, x_n)$$

o sea (como $\frac{\partial u_i}{\partial t} = x_i$)

$$x_1\frac{\partial f}{\partial u_1} + x_2\frac{\partial f}{\partial u_2} + \cdots + x_n\frac{\partial f}{\partial u_n} = \alpha t^{\alpha-1} f(x_1, x_2, \ldots, x_n)$$

Poniendo $t = 1$ obtenemos finalmente

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n} = \alpha f(x_1, x_2, \ldots, x_n)$$

(donde volvemos a llamar $x_1, x_2, \ldots, x_n$ a las variables de la función f. Nótese que estas derivadas quedan ya calculadas en $(x_1, x_2, \ldots, x_n)$, como tenía que ocurrir), que es relación que queríamos establecer. ■

Ejemplo 10. (El recíproco del Teorema de Euler sobre funciones homogéneas). Supongamos ahora que la función $f: \mathbb{R}^n \to \mathbb{R}$ satisface la relación

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_n \frac{\partial f}{\partial x_n} = \alpha f$$

(con las derivadas y la función calculadas en $(x_1, x_2, \ldots, x_n)$). Veamos entonces que esta función debe ser homogénea de grado α. Consideremos la función $\varphi: \mathbb{R}^+ \to \mathbb{R}$ de la variable t dada por

$$\varphi(t) = \frac{f(tx_1, tx_2, \ldots, tx_n)}{t^\alpha}$$

Si derivamos esta función obtenemos

$$\varphi\prime(t) = \frac{t^\alpha \left(x_1 \frac{\partial f}{\partial u_1} + x_2 \frac{\partial f}{\partial u_2} + \cdots + x_n \frac{\partial f}{\partial u_n} \right) - \alpha t^{\alpha-1} f(tx_1, tx_2, \ldots, tx_n)}{t^{2\alpha}}$$

en que llamamos $u_i = tx_i$, $i = 1, 2, \ldots, n$. Podemos reescribir esta última expresión como

$$\varphi\prime(t) = \frac{t^{\alpha-1} \left(u_1 \frac{\partial f}{\partial u_1} + u_2 \frac{\partial f}{\partial u_2} + \cdots + u_n \frac{\partial f}{\partial u_n} \right) - \alpha t^{\alpha-1} f(tx_1, tx_2, \ldots, tx_n)}{t^{2\alpha}}$$

Atendiendo a la relación que satisface (por hipótesis) la función f tenemos

$$\varphi\prime(t) = \frac{t^{\alpha-1} (\alpha f(tx_1, tx_2, \ldots, tx_n)) - \alpha t^{\alpha-1} f(tx_1, tx_2, \ldots, tx_n)}{t^{2\alpha}} = 0$$

Así que $\varphi'(t) = 0 \; \forall t \in \mathbb{R}^+$. Como consecuencia del Teorema del Valor Medio, concluimos que φ debe ser una función constante. De hecho, como $\varphi(1) = f(x_1, x_2, \ldots, x_n)$, tenemos

$$\varphi(t) = \frac{f(tx_1, tx_2, \ldots, tx_n)}{t^\alpha} = f(x_1, x_2, \ldots, x_n), \quad \forall t > 0$$

o sea $f(tx_1, tx_2, \ldots, tx_n) = t^\alpha f(x_1, x_2, \ldots, x_n)$, $\forall t > 0$, lo que nos dice que f es una función homogénea de grado α, como queríamos demostrar. ■

Ejemplo 11. Consideremos la función $z = z(x, y)$ que se compone con las funciones $x = x(u, v)$, $y = y(u, v)$. Ya sabemos entonces que las derivadas parciales respecto de u y v de la función compuesta son

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}$$
$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

Queremos ahora obtener la derivada parcial de segundo orden $\frac{\partial^2 z}{\partial u^2}$. Es decir, queremos obtener

$$\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial u}\right) = \frac{\partial}{\partial u}\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}\right) = \frac{\partial}{\partial u}\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}\right) + \frac{\partial}{\partial u}\left(\frac{\partial z}{\partial y}\frac{\partial y}{\partial u}\right)$$

Debemos ser muy cuidadosos al calcular estas nuevas derivadas parciales indicadas. Es necesario reconocer adecuadamente las nuevas funciones (junto con sus variables) que vamos a derivar. Por ejemplo, el símbolo $\frac{\partial x}{\partial u}$ que aparece en la expresión anterior representa la derivada parcial de la función $x = x(u, v)$. Derivar esta nueva función respecto de u no representa problema alguno, pues esto es $\frac{\partial}{\partial u}\left(\frac{\partial x}{\partial u}\right) = \frac{\partial^2 x}{\partial u^2}$. Una situación similar se presenta con la función $\frac{\partial y}{\partial u}$. Un problema menos evidente se presenta con las expresiones $\frac{\partial z}{\partial x}$ y $\frac{\partial z}{\partial y}$. Estas representan las nuevas funciones que se obtienen al derivar respecto de x y y a la función dada $z = z(x, y)$. Sin embargo, no está claro (no se hace explícito en esta notación) que estas derivadas parciales están calculadas en $x = x(u, v)$ y $y = y(u, v)$, (sugerimos al lector que revise nuevamente la fórmula de las derivadas parciales de la función compuesta establecida en el Teorema de esta sección), es decir, que estas funciones son nuevas funciones compuestas

$$\frac{\partial z}{\partial x}(x(u, v), y(x, v)), \qquad \frac{\partial z}{\partial y}(x(u, v), y(u, v))$$

que queremos derivar respecto de u. El esquema que debemos tener siempre presente es el siguiente: si $\square$ representa una función de las variables x, y, en la que x, y son funciones de u y v, entonces

$$\frac{\partial}{\partial u}\square = \frac{\partial}{\partial x}\square\frac{\partial x}{\partial u} + \frac{\partial}{\partial y}\square\frac{\partial y}{\partial u}$$

Este hecho ya lo aplicamos al principio del problema con la función $z = z(x, y)$ y ahora lo tenemos que aplicar con las funciones $\frac{\partial z}{\partial x}$ y $\frac{\partial z}{\partial y}$. Así,

$$\begin{aligned}\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial x}\right) &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)\frac{\partial x}{\partial u} + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)\frac{\partial y}{\partial u}\\ &= \frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial u} + \frac{\partial^2 z}{\partial y \partial x}\frac{\partial y}{\partial u}\end{aligned}$$

y

$$\begin{aligned}\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial y}\right) &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)\frac{\partial x}{\partial u} + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right)\frac{\partial y}{\partial u}\\ &= \frac{\partial^2 z}{\partial x \partial y}\frac{\partial x}{\partial u} + \frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial u}\end{aligned}$$

Retomando nuestro problema original tenemos

$$\begin{aligned}\frac{\partial^2 z}{\partial u^2} &= \frac{\partial}{\partial u}\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}\right) + \frac{\partial}{\partial u}\left(\frac{\partial z}{\partial y}\frac{\partial y}{\partial u}\right) \xrightarrow[\text{de dos funciones}]{\text{derivadas de productos}}\\ &= \frac{\partial z}{\partial x}\frac{\partial}{\partial u}\left(\frac{\partial x}{\partial u}\right) + \frac{\partial x}{\partial u}\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial x}\right) + \frac{\partial z}{\partial y}\frac{\partial}{\partial u}\left(\frac{\partial y}{\partial u}\right) + \frac{\partial y}{\partial u}\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial y}\right)\\ &= \frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial u^2} + \frac{\partial x}{\partial u}\left(\frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial u} + \frac{\partial^2 z}{\partial y \partial x}\frac{\partial y}{\partial u}\right) + \frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial u^2} + \frac{\partial y}{\partial u}\left(\frac{\partial^2 z}{\partial x \partial y}\frac{\partial x}{\partial u} + \frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial u}\right)\end{aligned}$$

o bien, simplificando (y suponiendo que la función z es lo suficientemente bien portada para que sus derivadas cruzadas sean iguales) nos queda

$$\frac{\partial^2 z}{\partial u^2} = \frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial u^2} + \frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial u}\right)^2 + \frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial u^2} + \frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial u}\right)^2 + 2\frac{\partial^2 z}{\partial x \partial y}\frac{\partial x}{\partial u}\frac{\partial y}{\partial u}$$

De manera análoga, el lector puede obtener las expresiones

$$\frac{\partial^2 z}{\partial v^2} = \frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial v^2} + \frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial v}\right)^2 + \frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial v^2} + \frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial v}\right)^2 + 2\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial v}\frac{\partial y}{\partial v}$$

$$\frac{\partial^2 z}{\partial u\partial v} = \frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial u\partial v} + \frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial x\partial y} + \frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + \frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial v}\frac{\partial y}{\partial y} + \frac{\partial^2 z}{\partial y\partial x}\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} \quad \blacksquare$$

Ejemplo 12. Obtengamos las derivadas de segundo orden de la función compuesta $z = f\left(u^2 + v^2, \frac{u}{v}\right)$. Llamando $x = u^2 + v^2$, $y = \frac{u}{v}$, tenemos que las derivadas de primer orden son

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} = 2u\frac{\partial z}{\partial x} + \frac{1}{v}\frac{\partial z}{\partial y}$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v} = 2v\frac{\partial z}{\partial x} - \frac{u}{v^2}\frac{\partial z}{\partial y}$$

Obtengamos $\frac{\partial^2 z}{\partial u^2}$ siguiendo las indicaciones del ejemplo anterior

$$\begin{aligned}
\frac{\partial^2 z}{\partial u^2} &= \frac{\partial}{\partial u}\left(\frac{\partial z}{\partial u}\right) = \frac{\partial}{\partial u}\left(2u\frac{\partial z}{\partial x} + \frac{1}{v}\frac{\partial z}{\partial y}\right) \\
&= 2u\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial x}\right) + \frac{\partial z}{\partial x}\frac{\partial}{\partial u}(2u) + \frac{1}{v}\frac{\partial}{\partial u}\left(\frac{\partial z}{\partial y}\right) + \frac{\partial z}{\partial y}\frac{\partial}{\partial u}\left(\frac{1}{v}\right) = \\
&= 2u\left(\frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial u} + \frac{\partial^2 z}{\partial y\partial x}\frac{\partial y}{\partial u}\right) + 2\frac{\partial z}{\partial x} + \frac{1}{v}\left(\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial u} + \frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial u}\right) \\
&= 2u\left(\frac{\partial^2 z}{\partial x^2}2u + \frac{\partial^2 z}{\partial y\partial x}\left(\frac{1}{v}\right)\right) + 2\frac{\partial z}{\partial x} + \frac{1}{v}\left(\frac{\partial^2 z}{\partial x\partial y}(2u) + \frac{\partial^2 z}{\partial y^2}\left(\frac{1}{v}\right)\right) \\
&= 4u^2\frac{\partial^2 z}{\partial x^2} + \frac{4u}{v}\frac{\partial^2 z}{\partial x\partial y} + \frac{1}{v^2}\frac{\partial^2 z}{\partial y^2} + 2\frac{\partial z}{\partial x}
\end{aligned}$$

Obtengamos $\frac{\partial^2 z}{\partial u\partial v}$

$$\begin{aligned}
\frac{\partial^2 z}{\partial u\partial v} &= \frac{\partial}{\partial v}\left(\frac{\partial z}{\partial u}\right) = \frac{\partial}{\partial v}\left(2u\frac{\partial z}{\partial x} + \frac{1}{v}\frac{\partial z}{\partial y}\right) \\
&= 2u\frac{\partial}{\partial v}\left(\frac{\partial z}{\partial x}\right) + \frac{\partial z}{\partial x}\frac{\partial}{\partial v}(2u) + \frac{1}{v}\frac{\partial}{\partial v}\left(\frac{\partial z}{\partial y}\right) + \frac{\partial z}{\partial y}\frac{\partial}{\partial v}\left(\frac{1}{v}\right) \\
&= 2u\left(\frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial v} + \frac{\partial^2 z}{\partial y\partial x}\frac{\partial y}{\partial v}\right) + \frac{1}{v}\left(\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial v} + \frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial v}\right) + \frac{\partial z}{\partial y}\left(-\frac{1}{v^2}\right) \\
&= 2u\left(\frac{\partial^2 z}{\partial x^2}2v + \frac{\partial^2 z}{\partial y\partial x}\left(-\frac{u}{v^2}\right)\right) + \frac{1}{v}\left(\frac{\partial^2 z}{\partial x\partial y}2v + \frac{\partial^2 z}{\partial y^2}\left(-\frac{u}{v^2}\right)\right) + \frac{\partial z}{\partial y}\left(-\frac{1}{v^2}\right) \\
&= 4uv\frac{\partial^2 z}{\partial x^2} - \frac{2u^2}{v^2}\frac{\partial^2 z}{\partial y\partial x} + 2\frac{\partial^2 z}{\partial x\partial y} - \frac{u}{v^3}\frac{\partial^2 z}{\partial y^2} - \frac{1}{v^2}\frac{\partial z}{\partial y}
\end{aligned}$$

Por último, obtengamos $\frac{\partial^2 z}{\partial v^2}$

$$\frac{\partial^2 z}{\partial v^2} = \frac{\partial}{\partial v}\left(\frac{\partial z}{\partial v}\right) = \frac{\partial}{\partial v}\left(2v\frac{\partial z}{\partial x} - \frac{u}{v^2}\frac{\partial z}{\partial y}\right)$$

$$= 2v\frac{\partial}{\partial v}\left(\frac{\partial z}{\partial x}\right) + \frac{\partial z}{\partial x}\frac{\partial}{\partial v}(2v) - \frac{u}{v^2}\frac{\partial}{\partial v}\left(\frac{\partial z}{\partial y}\right) - \frac{\partial z}{\partial y}\frac{\partial}{\partial v}\left(\frac{u}{v^2}\right)$$

$$= 2v\left(\frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial v} + \frac{\partial^2 z}{\partial y \partial x}\frac{\partial y}{\partial v}\right) + 2\frac{\partial z}{\partial x} - \frac{u}{v^2}\left(\frac{\partial^2 z}{\partial x \partial y}\frac{\partial x}{\partial v} + \frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial v}\right) - \frac{\partial z}{\partial y} - \left(\frac{2u}{v^3}\right)$$

$$= 2v\left(\frac{\partial^2 z}{\partial x^2}(2v) + \frac{\partial^2 z}{\partial y \partial x}\left(-\frac{u}{v^2}\right)\right) + 2\frac{\partial z}{\partial x} - \frac{u}{v^2}\left(\frac{\partial^2 z}{\partial x \partial y}(2v) + \frac{\partial^2 z}{\partial x^2}\left(-\frac{u}{v^2}\right)\right) + \frac{2u}{v^3}\frac{\partial z}{\partial y}$$

$$= 4v^2\frac{\partial^2 z}{\partial x^2} + \frac{u^2}{v^4}\frac{\partial^2 z}{\partial y^2} + 2\frac{\partial z}{\partial x} + \frac{2u}{v^3}\frac{\partial z}{\partial y} - \frac{4u}{v}\frac{\partial^2 z}{\partial x \partial y}$$ ■

Ejemplo 13. Una ecuación muy importante en la Física Matemática es la Ecuación de Laplace

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$

Supongamos que la función $z = f(x, y)$ satisface la Ecuación de Laplace. Veamos entonces que la función $z = f(x - 2y, 2x + y)$ también la satisface. En efecto, llamando $u = x - 2y$, $v = 2x + y$, tenemos

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial z}{\partial u} + 2\frac{\partial z}{\partial v}$$

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial^2 z}{\partial u^2}\frac{\partial u}{\partial x} + \frac{\partial^2 z}{\partial v \partial u}\frac{\partial v}{\partial x} + 2\left(\frac{\partial^2 z}{\partial u \partial v}\frac{\partial u}{\partial x} + \frac{\partial^2 z}{\partial v^2}\frac{\partial v}{\partial x}\right)$$

$$= \frac{\partial^2 z}{\partial u^2} + 4\frac{\partial^2 z}{\partial u \partial v} + 4\frac{\partial^2 z}{\partial v^2}$$

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial y} = \frac{\partial z}{\partial u}(-2) + \frac{\partial z}{\partial v}(1) = -2\frac{\partial z}{\partial u} + \frac{\partial z}{\partial v}$$

$$\frac{\partial^2 z}{\partial y^2} = -2\left(\frac{\partial^2 z}{\partial u^2}\frac{\partial u}{\partial y} + \frac{\partial^2 z}{\partial v \partial u}\frac{\partial v}{\partial y}\right) + \frac{\partial^2 z}{\partial u \partial v}\frac{\partial u}{\partial y} + \frac{\partial^2 z}{\partial v^2}\frac{\partial v}{\partial y}$$

$$= 4\frac{\partial^2 z}{\partial u^2} - 4\frac{\partial^2 z}{\partial u \partial v} + \frac{\partial^2 z}{\partial v^2}$$

Entonces

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 5\frac{\partial^2 z}{\partial u^2} + 5\frac{\partial^2 z}{\partial v^2} = 5\left(\frac{\partial^2 z}{\partial u^2} + \frac{\partial^2 z}{\partial v^2}\right) = 0$$

pues, por la hipótesis, la función original z satisface la Ecuación de Laplace. ■

Ejemplo 14. Una situación que se presenta con cierta frecuencia en las aplicaciones, es efectuar "cambios de variable" en expresiones diferenciales (expresiones que contienen una función determinada —de una o varias variables— y sus derivadas). Esto se logra aplicando correctamente la regla de la cadena en las expresiones de cambios de variables. Veamos un ejemplo importante.

Dada la función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ (suficientemente diferenciable), se le asocia a ella su *laplaciano*, denotado por $\nabla^2 f$ y definido como

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial z^2}$$

(Obsérvese que la Ecuación de Laplace del ejemplo anterior es $\nabla^2 z = 0$). Se pretende ver cuál es la expresión del laplaciano en coordenadas polares. Más bien, supongamos que en lugar de las coordenadas cartesianas x, y, con las cuales el laplaciano tiene el aspecto mostrado arriba, trabajamos con las coordenadas polares r, θ y con la función $z = F(r, \theta) = f(r\cos\theta, r\operatorname{sen}\theta)$. La pregunta es: ¿cuál es el aspecto del laplaciano con esta nueva función? (tenga en cuenta que ésta no es más que la composición de la función dada $z = f(x, y)$ con la función $G(r, \theta) = (r\cos\theta, r\operatorname{sen}\theta)$, que establece el cambio de coordenadas de polares a cartesianas). Esquemáticamente

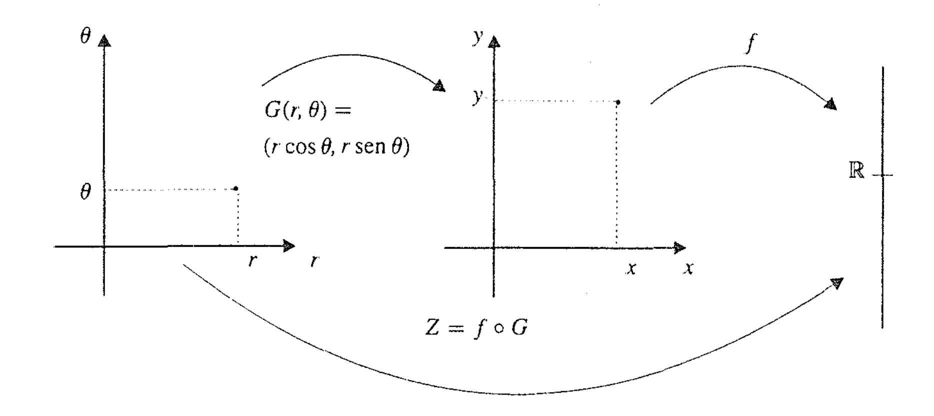


Figura 4. Esquema del paso de coordenadas cartesianas a polares.

De la función $z = f(r\cos\theta, r\operatorname{sen}\theta)$ tenemos

$$\frac{\partial z}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \operatorname{sen}\theta\frac{\partial f}{\partial y}$$

$$\frac{\partial z}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} = -r\operatorname{sen}\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}$$

Obtengamos las segundas derivadas $\dfrac{\partial^2 z}{\partial r^2}$ y $\dfrac{\partial^2 z}{\partial \theta^2}$

$$\begin{aligned}
\frac{\partial^2 z}{\partial r^2} &= \frac{\partial}{\partial r}\left(\frac{\partial z}{\partial r}\right) = \\
&= \cos\theta\left(\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y \partial x}\frac{\partial y}{\partial r}\right) + \operatorname{sen}\theta\left(\frac{\partial^2 f}{\partial x \partial y}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial r}\right) \\
&= \cos\theta\left(\frac{\partial^2 f}{\partial x^2}\cos\theta + \frac{\partial^2 f}{\partial x \partial y}\operatorname{sen}\theta\right) + \operatorname{sen}\theta\left(\frac{\partial^2 f}{\partial y \partial x}\cos\theta + \frac{\partial^2 f}{\partial y^2}\operatorname{sen}\theta\right) \\
&= \cos^2\theta\frac{\partial^2 f}{\partial x^2} + 2\operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial x \partial y} + \operatorname{sen}^2\theta\frac{\partial^2 f}{\partial y^2}
\end{aligned}$$

También

$$\frac{\partial^2 z}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left(\frac{\partial z}{\partial \theta}\right) = -r \operatorname{sen}\theta \frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) + \frac{\partial f}{\partial x}\frac{\partial}{\partial \theta}\left(-r \operatorname{sen}\theta\right)$$

$$+ r\cos\theta \frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right) + \frac{\partial f}{\partial y}\frac{\partial}{\partial \theta}\left(r\cos\theta\right)$$

$$= -r \operatorname{sen}\theta\left(\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial \theta} + \frac{\partial^2 f}{\partial y \partial x}\frac{\partial y}{\partial \theta}\right) + \frac{\partial f}{\partial x}(-r\cos\theta)$$

$$+ r\cos\theta\left(\frac{\partial^2 f}{\partial x \partial y}\frac{\partial x}{\partial \theta} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial \theta}\right) + \frac{\partial f}{\partial y}(-r \operatorname{sen}\theta)$$

$$= -r \operatorname{sen}\theta\left(\frac{\partial^2 f}{\partial x^2}(-r \operatorname{sen}\theta) + \frac{\partial^2 f}{\partial y \partial x}(r\cos\theta)\right) - r\cos\theta\frac{\partial f}{\partial x}$$

$$+ r\cos\theta\left(\frac{\partial^2 f}{\partial x \partial y}(-r \operatorname{sen}\theta) + \frac{\partial^2 f}{\partial y^2}(r\cos\theta)\right) - r \operatorname{sen}\theta\frac{\partial f}{\partial y}$$

$$= r^2 \operatorname{sen}^2\theta\frac{\partial^2 f}{\partial x^2} - 2r^2 \operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial x \partial y}$$

$$+ r^2\cos^2\theta\frac{\partial^2 f}{\partial y^2} - r\cos\theta\frac{\partial f}{\partial x} - r \operatorname{sen}\theta\frac{\partial f}{\partial y}$$

Buscamos despejar de las expresiones arriba obtenidas, el laplaciano $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$. Se observa que dividiendo por r^2 la derivada $\frac{\partial^2 z}{\partial \theta^2}$ y sumándola con $\frac{\partial^2 r}{\partial r^2}$ se llega a la expresión buscada, más un par de términos fáciles de identificar.

$$\frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} + \frac{\partial^2 z}{\partial r^2} = \operatorname{sen}^2\theta\frac{\partial^2 f}{\partial x^2} - 2\operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial x \partial y} + \cos^2\theta\frac{\partial^2 f}{\partial y^2}$$

$$- \frac{1}{r}\cos\theta\frac{\partial f}{\partial x} - \frac{1}{r}\operatorname{sen}\theta\frac{\partial f}{\partial y} + \cos^2\theta\frac{\partial^2 f}{\partial x^2}$$

$$+ 2\operatorname{sen}\theta\cos\theta\frac{\partial^2 f}{\partial x \partial y} + \operatorname{sen}^2\theta\frac{\partial^2 f}{\partial y^2}$$

$$= \underbrace{\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}}_{\text{Laplaciano}} - \frac{1}{r}\underbrace{\left(\cos\theta\frac{\partial f}{\partial x} + \operatorname{sen}\theta\frac{\partial f}{\partial y}\right)}_{\frac{\partial z}{\partial r}}$$

Entonces

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} + \frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r}$$ ■

Dejemos por un momento "las cuentas" y regresemos al "detalle" de la teoría, que nos permitirá comprender de una manera global el teorema demostrado en esta sección.

Ejercicios (Capítulo 3, Sección 2)

En los ejercicios 1–15, identifique la(s) funcion(es) g que se está(n) componiendo con la función f (la cual suponemos diferenciable), y obtenga las derivadas parciales de la función F indicada, las cuales involucran a la composición de f con g (ver ejemplo 6).

1. $F(x, y) = f(ax + by)$

2. $F(x, y) = xf(xy)$

3. $F(x, y) = x^2 f(x \operatorname{sen} y) - yf(2xy)$

4. $F(x, y) = f(2, xy)$

5. $F(x, y) = f(y, x)$

6. $F(x, y) = f(x, y) + f(y^2, x^2)$

7. $F(x, y) = f(x - y, x + y) + f(x + y, x - y)$

8. $F(x, y, z) = f(x + y, y - z)$

9. $F(x, y, z) = f(x, xy, xyz) + f(xyz, xy, x)$

10. $F(x, y, z) = f(xy^2z^3, x^3y^2z, x^2y^2z^2)$

11. $F(x, y, z) = f(x + 3y, 2y - 3z, 2x + 7y - 6z, x - y - z)$

12. $F(x, y, z) = f(1 - x, 2 - y, 3 - z)$

13. $F(x, y, z, u) = f(-x, -y, -z, -u)$

14. $F(x, y, z, u) = f(\operatorname{sen} x, \cos y, \tan z, \cot u)$

15. $F(x_1, x_2, x_3, x_4) = f(x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4)$

En los ejercicios 16–20 se dan funciones $F(t)$ que involucran la composición de una función real diferenciable f definida en algún espacio $\mathbb{R}^n$ y una función g diferenciable definida en $\mathbb{R}$ y que toma valores en $\mathbb{R}^n$. En cada caso identifique la función g y obtenga la derivada $F'(t)$.

16. $F(t) = f(8t^2 + 13t - 1)$

17. $F(t) = f(3t + 2, 5t - 4)$

18. $F(t) = (t^3 + 2)f(1, t)$

19. $F(t) = tf(t, t, t)$

20. $F(t) = (\operatorname{sen}^2 t^3) \cos f(\operatorname{sen} t, \cos t)$

21. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable. Considere la función compuesta $F(x, y) = f(f(x, y), f(x, y))$. Aplicando la regla de la cadena, obtenemos que la derivada parcial de F respecto de x es

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x}\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x}\right)^2 + \frac{\partial f}{\partial y}\frac{\partial f}{\partial x}$$

Por ejemplo, si consideramos la función $f(x, y) = 7x - 9y$, tenemos que $F(x, y) = f(7x - 9y, 7x - 9y) = 7(7x - 9y) - 9(7x - 9y) = -14x + 18y$, de donde $\frac{\partial F}{\partial x} = -14$. Por otra

parte, $\left(\frac{\partial f}{\partial x}\right)^2 + \frac{\partial f}{\partial y}\frac{\partial f}{\partial x} = (7)^2 + (-9)(7) = -14 = \frac{\partial F}{\partial x}$. Sin embargo, si tomamos la función $f(x, y) = x^2 + 3xy$, tenemos $F(x, y) = f(x^2 + 3xy, x^2 + 3xy) = (x^2 + 3xy)^2 + 3(x^2 + 3xy)(x^2 + 3xy) = 4(x^2 + 3xy)^2$, de donde $\frac{\partial F}{\partial x} = 8(x^2 + 3xy)(2x + 3y)$, y, por otra parte se tiene $\left(\frac{\partial f}{\partial x}\right)^2 + \frac{\partial f}{\partial y}\frac{\partial f}{\partial x} = (2x + 3y)^2 + (3x)(2x + 3y) = (2x + 3y)(5x + 3y)$, expresión evidentemente diferente de $\frac{\partial F}{\partial x}$. Explique dónde está el error de esta (aparente) contradicción.

22. Considere la función $z = z(x, y)$. Al componer ésta con la función $f(x, y) = (x + y, x + 3)$ se obtiene $z = z(x + y, x + 3)$. Aplicando la regla de la cadena, obtenemos que la derivada parcial de esta última función respecto de x es

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial x}\frac{\partial}{\partial x}(x + y) + \frac{\partial z}{\partial y}\frac{\partial}{\partial x}(x + 3) = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}$$

de donde, eliminando $\frac{\partial z}{\partial x}$, se obtiene que $\frac{\partial z}{\partial y} = 0$, lo cual, evidentemente es (en general) falso. Explique el error en este razonamiento.

23. Sea $F(t) = f(t \operatorname{sen} t, t, t^2)$, donde f es diferenciable. Suponga que grad $f(0, 0, 0) = (2, 4, 7)$. Halle $F'(0)$.

24. Sea $g: \mathbb{R}^2 \to \mathbb{R}^3$, $g = (g_1, g_2, g_3)$ una función diferenciable tal que grad $g_1(0, 0) = (1, 2)$, grad $g_2(0, 0) = (0, 10)$, grad $g_3(0, 0) = (3, 1)$, y sea $f: \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable tal que grad $f(g(0, 0)) = (3, -4, 2)$. Hallar grad$(f \circ g)(0, 0)$.

25. Sea $F: \mathbb{R}^n \to \mathbb{R}$ la función definida como

$$F(x_1, x_2, \ldots, x_n) = f(x_1, x_1x_2, x_1x_2x_3, \ldots, x_1x_2 \ldots x_n)$$

donde $f: \mathbb{R}^n \to \mathbb{R}$ es una función diferenciable. Si grad $f(1, 1, \ldots, 1) = (a_1, a_2, \ldots, a_n)$, halle grad $F(1, 1, \ldots, 1)$.

26. Sea $F: \mathbb{R}^n \to \mathbb{R}$ la función definida como

$$F(x_1, x_2, \ldots, x_n) = f(x_1, x_1 + x_2, \ldots, x_1 + x_2 + \cdots + x_n)$$

donde $f: \mathbb{R}^n \to \mathbb{R}$ es una función diferenciable. Si grad $f(1, 2, \ldots, n) = (1, 2, \ldots, n)$, halle grad $F(1, 1, \ldots, 1)$.

27. Sea $F(x, y) = f(x + 3y, 2x - y)$, donde $f: \mathbb{R}^2 \to \mathbb{R}$ es diferenciable. Suponga que grad $f(0, 0) = (4, -3)$. Determine la derivada de la función F en el origen en la dirección del vector $\mathbf{v} = (1, 1)$.

28. Sea $F(x, y) = f(x^2 + y^3, 5x + 7y, 3x^2y, x^3y^7)$, donde $f: \mathbb{R}^4 \to \mathbb{R}$ es diferenciable. Suponga que grad $f(0, 0, 0, 0) = (a, b, c, d)$. Determine la derivada de la función F en el origen en la dirección del vector unitario $\mathbf{v} = (\alpha_1, \alpha_2)$.

29. Sea $F(x, y) = f(x^2 + y, 3xy)$, donde $f: \mathbb{R}^2 \to \mathbb{R}$ es diferenciable. Suponga que grad $f(2, 3) = (5, 4)$. Hallar la dirección del mayor crecimiento de la función F en el punto $(1, 1)$.

30. Sean $f, g, h: \mathbb{R} \to \mathbb{R}$ tres funciones diferenciables cuyas gráficas pasan por el origen de coordenadas, el cual es un punto crítico para cada una de ellas. Demuestre que el plano tangente a la gráfica de la función $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = f(g(x) + h(y))$ es el plano $z = 0$. ¿Ocurre lo mismo con la función $F(x, y) = f(g(x)h(y))$?, ¿y con la función $F(x, y) = f(g(x), h(y))$?

31. Sea $F(x, y) = f(f(x, y), f(x, y))$, donde $f: \mathbb{R}^2 \to \mathbb{R}$ es una función diferenciable tal que $f(0, 0) = 0$, grad $f(0, 0) = (1, 2)$. Hallar la ecuación del plano tangente a la gráfica de F en el origen.

32. **a.** Considere las funciones $\phi, \psi: \mathbb{R}^3 \to \mathbb{R}$, $\phi(x, y, z) = x^3 - 3xyz^2$, $\psi(x, y, z) = xy + xy^2 + z^4$. Determine la ecuación del plano Π en $\mathbb{R}^3$ en el que se encuentran los vectores $\mathbf{v}_1 = \operatorname{grad}\phi(1, 1, 1)$, $\mathbf{v}_2 = \operatorname{grad}\psi(1, 1, 1)$.

b. Sea $F: \mathbb{R}^2 \to \mathbb{R}$ la función $F(x, y) = 5x + 4y - 3xy$. Determine la función $G: \mathbb{R}^3 \to \mathbb{R}$, $G(x, y, z) = F(\phi(x, y, z), \psi(x, y, z))$. Demuestre que el vector grad $G(1, 1, 1)$ se encuentra en el plano Π.

33. Demuestre que el resultado del ejercicio anterior es un hecho general: si $G: \mathbb{R}^3 \to \mathbb{R}$ es la función $G(x, y, z) = F(\phi(x, y, z), \psi(x, y, z))$, en que $F: \mathbb{R}^2 \to \mathbb{R}$, $\phi, \psi: \mathbb{R}^3 \to \mathbb{R}$ son funciones diferenciables, entonces los vectores grad $\phi(\mathbf{p})$, grad $\psi(\mathbf{p})$ y grad $G(\mathbf{p})$, donde $\mathbf{p}$ es un punto dado de $\mathbb{R}^3$, se encuentran en un mismo plano.

34. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable tal que $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 0$, $f(x, y) \neq 0\ \forall (x, y) \in \mathbb{R}^2$. Considere la función $F: \mathbb{R}^2 \to \mathbb{R}$ dada por $F(x, y) = \frac{x}{f(x,y)}$. Demuestre que $\frac{\partial F}{\partial x}(x, y) + \frac{\partial F}{\partial y}(x, y) = \frac{F(x,y)}{x}$.

35. Demuestre que la función $F(x, y) = f\left(\frac{y}{x^2 - y^2}\right)$ donde f es una función real diferenciable de variable real, satisface la ecuación diferencial parcial $(x^2 + y^2)\frac{\partial F}{\partial x} + 2xy\frac{\partial F}{\partial y} = 0$.

36. Dada la función $f: \mathbb{R} \to \mathbb{R}$ tres veces diferenciable, se define la *derivada de Schwarz de f*, denotada por Sf, como

$$Sf(t) = \frac{f'''(t)}{f'(t)} - \frac{3}{2}\left(\frac{f''(t)}{f'(t)}\right)^2$$

a. Sea $g: \mathbb{R} \to \mathbb{R}$ una función tres veces diferenciable. Considere la composición $g \circ f: \mathbb{R} \to \mathbb{R}$. Demuestre que $S(g \circ f)(t) = (Sg(f(t)))(f'(t))^2 + Sf(t)$.

b. Sea $F(t) = \frac{af(t)+b}{cf(t)+d}$, donde $ad - bc \neq 0$. Demuestre que $SF(t) = Sf(t)$. Concluya que $S\big((1/f)(t)\big) = Sf(t)$.

En los ejercicios 37–48 se da una función F que involucra una composición de una cierta función (suficientemente diferenciable) f con otra cierta función g. Obtenga las derivadas parciales de segundo orden de la función F

37. $F(x, y) = f(ax + by)$

38. $F(x, y) = f(ax + by, cx + dy)$

39. $F(x, y) = xf(y, y) + yf(x, x)$

40. $F(x, y) = (x^2 + y^2)f(x, 2y)$

41. $F(x, y) = f(x, xy, x + y)$

42. $F(x, y) = \int_x^{f(x,y)} g(t)dt$, donde $g: \mathbb{R} \to \mathbb{R}$ es una función continua.

43. $F(x, y) = \int_{xy}^{f(x+y, x-y)} g(t)dt$

44. $F(x, y) = \int_{yf(\operatorname{sen} x, \cos y, xy)}^{x} g(t)\, dt$

45. $F(x, y, z) = f(ax + by + cz)$

46. $F(x, y, z) = xf(xy, xz)$

47. $F(x, y, z) = f(xyz, xy, x)$

48. $F(x, y, z) = f(f(x, y, z), y, z)$

49. Obtenga las derivadas parciales de tercer orden de la función $F(x, y) = f(ax + by, cx + dy)$, en donde $f: \mathbb{R}^2 \to \mathbb{R}$ es una función de clase $\mathscr{C}^3$.

(*) **50.** Sea $F: \mathbb{R}^n \to \mathbb{R}$ la función dada por

$$F(x_1, x_2, \ldots, x_n) = f(x_n, x_1, x_2, \ldots, x_{n-1})$$

donde $f: \mathbb{R}^n \to \mathbb{R}$ es una función de clase $\mathscr{C}^2$. Obtenga las derivadas parciales de segundo orden de la función F.

51. Sea $F(x, y) = x^\alpha \phi\left(\frac{y}{x}\right)$, en que ϕ es una función real dos veces diferenciable, de una variable real, y α es un número real. Demuestre que

a. $x\frac{\partial F}{\partial x}(x, y) + y\frac{\partial F}{\partial y}(x, y) = \alpha F(x, y)$

b. $x^2\frac{\partial^2 F}{\partial x^2}(x, y) + 2xy\frac{\partial^2 F}{\partial x \partial y}(x, y) + y^2\frac{\partial^2 F}{\partial y^2}(x, y) = \alpha(\alpha - 1)F(x, y)$

52. Sea $F(x, y, z) = x^\alpha \phi\left(\frac{y}{x}, \frac{z}{x}\right)$, donde ϕ es una función real de dos variables, de clase $\mathscr{C}^2$, y α es un número real. Demuestre que

a. $x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} + z\frac{\partial F}{\partial z} = \alpha F$

b. $x^2\frac{\partial^2 F}{\partial x^2} + y^2\frac{\partial^2 F}{\partial y^2} + z^2\frac{\partial^2 F}{\partial z^2} + 2xy\frac{\partial^2 F}{\partial x \partial y} + 2xz\frac{\partial^2 F}{\partial x \partial z} + 2yz\frac{\partial^2 F}{\partial y \partial z} = \alpha(\alpha - 1)F$

53. Concilie los resultados de los dos ejercicios anteriores con lo establecido en el teorema de Euler sobre funciones homogéneas estudiado en el apéndice II de la sección 12 del capítulo 2.

54. Sea $f: \mathbb{R}^n \to \mathbb{R}$ una función diferenciable homogénea de grado α (es decir, se tiene que $f(tx_1, tx_2, \ldots, tx_n) = t^\alpha f(x_1, x_2, \ldots, x_n)$, $\forall t > 0$, $\forall (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$). Demuestre que esta función se puede escribir (para x_1 positivo) como

$$f(x_1, x_2, \ldots, x_n) = x_1^\alpha \phi\left(\frac{x_2}{x_1}, \frac{x_3}{x_1}, \ldots, \frac{x_n}{x_1}\right)$$

donde ϕ es una función real de $n - 1$ variables. Use este hecho para verificar directamente que

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n} = \alpha f$$

Obsérvese que el inciso **a** de los ejercicios 51 y 52 son casos particulares de este ejercicio.

55. (Las funciones que aparecen en este ejercicio se consideran de clase $\mathscr{C}^2$). Se dice que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ satisface la *ecuación de Laplace* si

$$L(f(x, y)) = \frac{\partial^2 f}{\partial x^2}(x, y) + \frac{\partial^2 f}{\partial y^2}(x, y) = 0 \qquad \forall (x, y) \in U$$

a. Sean $\phi, \psi: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ dos funciones tales que

$$\frac{\partial \phi}{\partial u} = \frac{\partial \psi}{\partial v} \qquad \frac{\partial \phi}{\partial v} = -\frac{\partial \psi}{\partial u}$$

Demuestre que ϕ y ψ satisfacen la ecuación de Laplace.

b. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función que satisface la ecuación de Laplace, y sean ϕ y ψ dos funciones como las del inciso anterior. Demuestre que la función $F(u, v) = f(\phi(u, v), \psi(u, v))$ satisface la ecuación de Laplace.

56. Con el resultado del ejercicio anterior, demuestre que si la función $f(x, y)$ satisface la ecuación de Laplace, entonces la función

$$F(x, y) = f\left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2}\right)$$

también la satisface.

57. Demuestre que si la función $f(x, y)$ satisface la ecuación de Laplace, entonces la función

$$F(x, y) = f\left(\frac{\pm x}{x^2 + y^2}, \frac{\pm y}{x^2 + y^2}\right)$$

también la satisface.

58. Considere la ecuación

$$\frac{\partial^2 f}{\partial x^2}(x, y) + 2xy^2 \frac{\partial f}{\partial x}(x, y) + 2(y - y^3)\frac{\partial f}{\partial y}(x, y) + x^2 y^2 f(x, y) = 0 \qquad \text{e}$$

donde $f: \mathbb{R}^2 \to \mathbb{R}$ es una función de clase $\mathscr{C}^2$. Diremos que (e) es la "ecuación asociada a la función $f(x, y)$". Demuestre que (e) es también la ecuación asociada a la función $F(u, v) = f(uv, 1/v)$. Es decir, (e) no cambia su forma si en lugar de $f(x, y)$ ponemos $F(u, v) = f(uv, 1/v)$.

59. Considere la ecuación diferencial ordinaria

$$ax^2 f''(x) + bxf'(x) + cf(x) = 0$$

en la que a, b, c son constantes (llamada "Ecuación de Euler-Cauchy"). Demuestre que si hacemos el cambio de variable independiente x por t, según la fórmula $x = e^t$, la ecuación se transforma en

$$\alpha F''(t) + \beta F'(t) + \gamma F(t) = 0$$

donde α, β, γ son constantes, y $F(t) = f(e^t)$. Generalice este resultado para ecuaciones del tipo

$$a_n x^n f^{(n)}(x) + a_{n-1} x^{n-1} f^{(n-1)}(x) + \cdots + a_1 x f'(x) + a_0 f(x) = 0$$

60. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathscr{C}^2$. Considere la función $F(u, v) = f(uv, \frac{1}{2}(u^2 - v^2))$. Demuestre que

$$(u^2 + v^2)\left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right] = \left(\frac{\partial F}{\partial u}\right)^2 + \left(\frac{\partial F}{\partial v}\right)^2$$

3.3 Regla de la cadena. Perspectiva general

Cuando se ve el teorema presentado en la sección anterior (regla de la cadena), una vez apreciada la sencillez del resultado análogo para funciones de una variable, el lector se puede desconcertar ante la (aparente) complejidad de las fórmulas de las derivadas parciales de la función compuesta ahí establecidas. En esta sección pretendemos ubicar al lector en el ángulo adecuado para que vea que la regla de la cadena en el caso más general (aun que en el teorema de la sección anterior) sigue siendo un resultado de asombrosa sencillez y elegancia, como lo es en el caso de funciones de una variable. Para lograr este objetivo tendremos que utilizar un lenguaje algebraico un poco más sofisticado que el que hemos venido usando: este es el precio que hay que pagar por ver de manera global a qué se refiere la regla de la cadena. Dicho lenguaje es el del álgebra lineal (matrices y transformaciones lineales). El lector que no esté familiarizado con él, debe estudiar primeramente la sección 7 del capítulo 1, en donde se hace una breve exposición de los conceptos algebraicos que se manejarán en las discusiones que se presentan en esta sección.

Recordemos primeramente lo que, con palabras, establece la regla de la cadena para funciones de una variable: "la composición de dos funciones diferenciables es diferenciable y su derivada es el producto de las derivadas de cada una de las funciones que se están componiendo". Uno de los objetivos de esta sección será convencernos de que un enunciado similar vale para el caso de funciones entre espacios de dimensiones mayores a uno. Llamamos, sin embargo, la atención al hecho de que hasta este momento no se ha definido el concepto de *derivada* de una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. Hemos definido a qué se refieren los términos "derivadas parciales", "derivadas direccionales", "gradiente" y "diferenciales" para funciones de varias variables, pero nunca hemos dicho qué es "la derivada" para este tipo de funciones y... éste es un buen momento para hacerlo.

Trabajaremos con funciones $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ definidas en el conjunto abierto U de $\mathbb{R}^n$ y tomando valores en $\mathbb{R}^m$. Entonces $f = (f_1, f_2, \ldots, f_m)$ donde $f_i: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, son las funciones coordenadas de f. Para definir la derivada de esta función en un punto $\mathbf{x}_0 \in U$, recordemos que en la sección 1 se había dicho que tal función es diferenciable en $\mathbf{x}_0$ (lo cual querrá decir que "la derivada" de la función existe en $\mathbf{x}_0$) si y sólo si todas y cada una de sus funciones componentes f_i lo eran. En este momento tendremos que establecer esta misma idea de una manera diferente, un poco más sofisticada, pero de mayor utilidad para nuestras intenciones.

Definición. Considere la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ definida en el conjunto abierto U de $\mathbb{R}^n$ y sea $\mathbf{x}_0 \in U$. Se dice que esta función es *diferenciable* en $\mathbf{x}_0$ si existe una transformación lineal $f'(\mathbf{x}_0): \mathbb{R}^n \to \mathbb{R}^m$, llamada *derivada* de f en $\mathbf{x}_0$, tal que $f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)\mathbf{h} + r(\mathbf{h})$, donde $\lim_{\mathbf{h}\to 0} \frac{r(\mathbf{h})}{\|\mathbf{h}\|}$ (para $\mathbf{h} \in \mathbb{R}^n$ tal que $\mathbf{x}_0 + \mathbf{h} \in U$). ■

NOTA: Siendo r una función que toma valores en $\mathbb{R}^m$, la interpretación del límite anterior sería de que todas sus funciones componentes tenderán a cero cuando, al dividirlas por $\|\mathbf{h}\|$, $\mathbf{h}$ tiende a cero.

En esta definición aparece ya el término "transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$", que se estudia en los cursos de Álgebra Lineal. Sabemos que estos objetos abstractos (las transformaciones lineales) tienen representaciones concretas por medio de matrices. De hecho, la transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ tiene asociada una matriz de orden $m \times n$ que la representa, la cual está construida de la forma siguiente: si $\beta_1 = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y $\beta_2 = \{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_m\}$ son las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, entonces el vector $T(\mathbf{e}_j)$ = imagen bajo T del j-ésimo vector de la base canónica de $\mathbb{R}^n$, se puede escribir como combinación lineal de los vectores de β_2; es decir, existen constante a_{ij} tales que

$$T(\mathbf{e}_j) = a_{1j}\bar{\mathbf{e}}_1 + a_{2j}\bar{\mathbf{e}}_2 + \cdots + a_{mj}\bar{\mathbf{e}}_m$$

A la matriz $A = (a_{ij})$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, cuya j-ésima columna está constituída por las coordenadas de la imagen del j-ésimo vector $\mathbf{e}_j$ de base canónica en $\mathbb{R}^n$ respecto de la base canónica de $\mathbb{R}^m$, se le llama matriz de transformación T respecto de las bases β_1 y β_2

Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable en $\mathbf{x}_0 \in U$. Consideremos el vector $\mathbf{h} = t\mathbf{e}_j$, $t \in \mathbb{R}$. Con t suficientemente pequeño podemos tener $\mathbf{x}_0 + \mathbf{h} \in U$. Entonces

$$\begin{aligned} f(\mathbf{x}_0 + t\mathbf{e}_j) &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(t\mathbf{e}_j) + r(\mathbf{h}) \\ &= f(\mathbf{x}_0) + tf'(\mathbf{x}_0)\mathbf{e}_j + r(\mathbf{h}) \end{aligned}$$

(la última igualdad usa la linealidad de $f'(\mathbf{x}_0)$). Se observa que $\|\mathbf{h}\| = \pm t$, y por lo tanto podemos escribir

$$f'(\mathbf{x}_0)\mathbf{e}_j = \frac{f(\mathbf{x}_0 + t\mathbf{e}_j) - f(\mathbf{x}_0)}{t} \mp \frac{r(\mathbf{h})}{\|\mathbf{h}\|}$$

Al tomar el límite cuando $t \to 0$ (i.e. cuando $\|\mathbf{h}\| \to 0$) el segundo sumando del lado derecho tiende a cero (por la definición de arriba) quedando sólo

$$f'(\mathbf{x}_0)\mathbf{e}_j = \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_j) - f(\mathbf{x}_0)}{t}$$

Al descomponer f como $(f_1, f_2, \ldots, f_n)$, y tomando el límite en cada una de las coordenadas vemos que éstas son las derivadas parciales $\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$, $i = 1, 2, \ldots, m$. Es decir

$$f'(\mathbf{x}_0)\mathbf{e}_j = \left(\frac{\partial f_1}{\partial x_j}(\mathbf{x}_0), \frac{\partial f_2}{\partial x_j}(\mathbf{x}_0), \ldots, \frac{\partial f_m}{\partial x_j}(\mathbf{x}_0) \right) = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)\bar{e}_i$$

Entonces la j-ésima columna de la matriz que representa a $f'(\mathbf{x}_0)$ está formada por las derivadas parciales de las funciones componentes respecto de su j-ésima variable. De aquí que la matriz que representa a la transformación $f'(\mathbf{x}_0): \mathbb{R}^n \to \mathbb{R}^m$ es

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

A esta matriz $m \times n$, en cuya i-ésima linea y j-ésima columna aparece la derivada parcial $\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$ de la i-ésima función componente de f respecto de su j-ésima variable, evaluada en $\mathbf{x}_0$, se le llama *matriz jacobiana* de la función f en $\mathbf{x}_0$ y se denota $Jf(\mathbf{x}_0)$. Esta es entonces *la derivada* de la función diferenciable f en $\mathbf{x}_0$ (identificando, como siempre, a la transformación lineal $f'(\mathbf{x}_0): \mathbb{R}^n \to \mathbb{R}^m$ con la matriz que la representa).

Veamos, por ejemplo, el caso en el que $m = 1$. En este caso la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$ será diferenciable en $\mathbf{x}_0 \in U$ si se da la transformación lineal $f'(\mathbf{x}_0): \mathbb{R}^n \to \mathbb{R}$ cuya representación matricial es la matriz $1 \times n$

$$Jf(\mathbf{x}_0) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \frac{\partial f}{\partial x_2}(\mathbf{x}_0) \ldots \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right]$$

tal que

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)\mathbf{h} + r(\mathbf{h}), \text{ con } \lim_{\mathbf{h}\to\mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Observe que si $\mathbf{h} = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ (tal que $\mathbf{x}_0 + \mathbf{h} \in U$) se tiene

$$f'(\mathbf{x}_0)\mathbf{h} = Jf(\mathbf{x}_0)\mathbf{h} = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

$$= \frac{\partial f}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial f}{\partial x_2}(\mathbf{x}_0)h_2 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0)h_n$$

Comparando esto con la definición que habíamos dado de diferenciabilidad para funciones de varias variables (capítulo 2 sección 6), vemos que se trata exactamente de la misma definición, sólo que en aquélla se hablaba de la existencia de las derivadas parciales y ahora hablamos de la transformación lineal $f'(\mathbf{x}_0)$ —la derivada de f en $\mathbf{x}_0$— o de modo más preciso, de la matriz que la representa (es decir, el residuo $r(\mathbf{h})$ queda determinado de la misma manera en ambas definiciones y se le pide que cumpla la misma propiedad).

Es posible demostrar (lo dejamos como ejercicio para el lector) que la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ será diferenciable en $\mathbf{x}_0 \in U$ (según la definición dada en esta sección) si y sólo si sus funciones coordenadas $f_i: U \subseteq \mathbb{R}^n \to \mathbb{R}$ lo son (con la definición de diferenciabilidad del capítulo 2). Más aún, pensando en que la derivada de la i-ésima función coordenada f_i en $\mathbf{x}_0$ es la matriz de orden $1 \times n$

$$Jf_i(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_i}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_i}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_i}{\partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

tenemos que la derivada de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es la matriz que en su i-ésima línea tiene la derivada de su i-ésima función componente. Esquematicamente

$$Jf(\mathbf{x}_0) = \begin{bmatrix} \leftarrow & Jf_1(\mathbf{x}_0) & \rightarrow \\ \leftarrow & Jf_2(\mathbf{x}_0) & \rightarrow \\ & \vdots & \\ \leftarrow & Jf_m(\mathbf{x}_0) & \rightarrow \end{bmatrix}$$

Obsérvese también que la matriz $1 \times n$, $Jf_i(\mathbf{x}_0)$, derivada de la función $f_i: U \subseteq \mathbb{R}^n \to \mathbb{R}$, se identifica de manera natural con el vector gradiente de f_i en $\mathbf{x}_0$ (poniendo comas a los espacios en la matriz, y cambiando los corchetes por paréntesis normales)

$$Jf_i(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_i}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_i}{\partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

$$\updownarrow \text{ identificación}$$

$$\operatorname{grad} f_i(\mathbf{x}_0) = \left(\frac{\partial f_i}{\partial x_1}(\mathbf{x}_0), \ldots, \frac{\partial f_i}{\partial x_n}(\mathbf{x}_0) \right)$$

teniéndose además que, bajo esta identificación, $Jf_i(\mathbf{x}_0)\mathbf{h} = \operatorname{grad} f(\mathbf{x}_0) \cdot \mathbf{h}$ (producto punto de vectores). Así, el vector gradiente de una función de n variables definido en el capítulo anterior, es "como la derivada de la función" en el sentido establecido anteriormente.

Veamos un par de ejemplos.

Ejemplo 1. La función $f: \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x, y) = (x^2 + 3y^2, 5x^3 + 2y^6)$ es diferenciable en todo el dominio y su derivada en un punto arbitrario (x, y) está dada por la matriz (aquí $f_1(x, y) = x^2 + 3y^2$, $f_2(x, y) = 5x^3 + 2y^6$

$$Jf = \begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 6y \\ 15x^2 & 12y^5 \end{bmatrix}$$ ■

Ejemplo 2. La función $f: \mathbb{R}^2 \to \mathbb{R}^3$ dada por $f(x, y) = (\text{sen}(x + y), xe^{x+y}, x + y)$ es diferenciable en todo su dominio (sus funciones coordenadas $f_1(x, y) = \text{sen}(x+y)$, $f_2(x, y) = xe^{x+y}$, $f_3(x, y) = x + y$, lo son) y su derivada en el punto $(0, 0)$ está dada por la matriz

$$Jf(0,0) = \begin{bmatrix} \dfrac{\partial f_1}{x}(0,0) & \dfrac{\partial f_1}{y}(0,0) \\ \dfrac{\partial f_2}{x}(0,0) & \dfrac{\partial f_2}{y}(0,0) \\ \dfrac{\partial f_3}{x}(0,0) & \dfrac{\partial f_3}{y}(0,0) \end{bmatrix} = \begin{bmatrix} \cos(x+y) & \cos(x+y) \\ e^{x+y}(x+1) & xe^{x+y} \\ 1 & 1 \end{bmatrix}_{\substack{x=0 \\ y=0}} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$ ■

Ahora estamos en posibilidades de establecer la regla de la cadena en el caso general. En principio trataremos de entender bien el resultado y dejaremos para el final el enunciado riguroso del teorema y su demostración. Consideremos entonces dos funciones diferenciables f y g, como en la figura 1

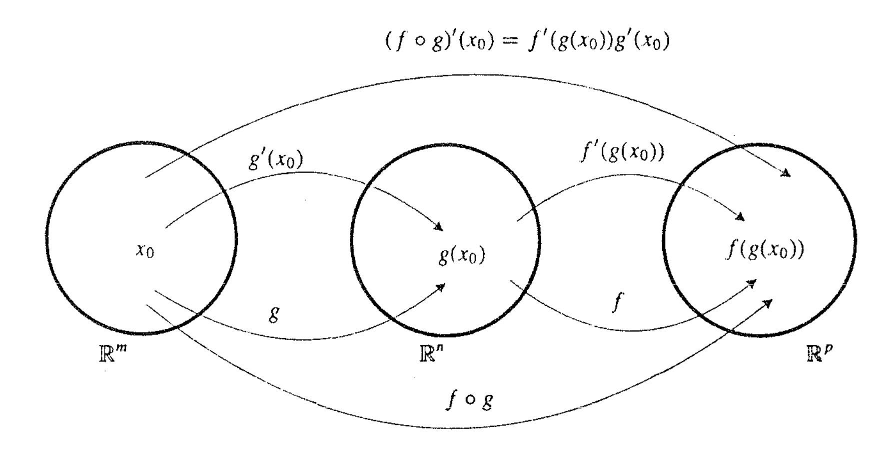


Figura 1. La composición de las funciones f y g.

Supongamos entonces que la función $g: \mathbb{R}^m \to \mathbb{R}^n$ es diferenciable en $\mathbf{x}_0$ y la función $f: \mathbb{R}^n \to \mathbb{R}^p$ es diferenciable en $g(\mathbf{x}_0)$. Lo que dice la regla de la cadena es que la función compuesta $f \circ g: \mathbb{R}^m \to \mathbb{R}^p$ será diferenciable en $\mathbf{x}_0$ y que

$$(f \circ g)'(\mathbf{x}_0) = f'(g(\mathbf{x}_0))g'(\mathbf{x}_0)$$

entendiéndose el lado derecho de esta expresión como una composición de transformaciones lineales (¡la misma fórmula que en el caso de una variable!), o bien, en términos matriciales

$$J(f \circ g)(\mathbf{x}_0) = Jf(g(\mathbf{x}_0))Jg(\mathbf{x}_0)$$

entendiéndose el lado derecho de esta última expresión como una multiplicación de matrices. Es decir, sigue siendo cierto, al igual que en el caso de funciones de una sola variable, que: *la (matriz que representa a la) derivada de la composición es igual al producto de las (matrices que representan a las) derivadas de las funciones componentes.*

Obsérvese que $Jg(\mathbf{x}_0)$ es una matriz $n \times m$ y $Jf(g(\mathbf{x}_0))$ una matriz $p \times n$, de modo que el producto $Jf(g(\mathbf{x}_0))Jg(\mathbf{x}_0)$ está bien determinado y da por resultado una matriz $p \times m$, como debe ser la derivada de $J(f \circ g)(\mathbf{x}_0)$.

Veamos nuevamente el caso $p = 1$ que ya hemos estudiado en la sección anterior. Tenemos entonces la función $g: \mathbb{R}^m \to \mathbb{R}^n$ diferenciable en $\mathbf{x}_0$ y la función $f: \mathbb{R}^n \to \mathbb{R}$ diferenciable en $g(\mathbf{x}_0)$. Si escribimos $g = (g_1, g_2, \ldots, g_n)$ donde $g_i: \mathbb{R}^m \to \mathbb{R}$, $i = 1, 2, \ldots, n$ son las funciones coordenadas de la función g, entonces la derivada de g en $\mathbf{x}_0$ es la matriz

$$Jg(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial g_1}{\partial x_m}(\mathbf{x}_0) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial g_2}{\partial x_m}(\mathbf{x}_0) \\ \vdots & & & \\ \frac{\partial g_n}{\partial x_1}(\mathbf{x}_0) & \frac{\partial g_n}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial g_n}{\partial x_m}(\mathbf{x}_0) \end{bmatrix}$$

en tanto que la derivada de $f: \mathbb{R}^n \to \mathbb{R}$ en $b = g(\mathbf{x}_0)$ es

$$Jf(b) = \begin{bmatrix} \frac{\partial f}{\partial y_1}(b) & \frac{\partial f}{\partial y_2}(b) & \cdots & \frac{\partial f}{\partial y_n}(b) \end{bmatrix}$$

(denotamos por $(y_1, y_2, \ldots, y_n)$ a las variables de la función f).

Entonces la derivada de la función compuesta $f \circ g$ en $\mathbf{x}_0$ debe ser la matriz $J(f \circ g)(\mathbf{x}_0)$ de orden $1 \times m$ que se obtiene como el producto de las matrices $Jf(b)$ y $Jg(\mathbf{x}_0)$. Es decir

$$\begin{aligned} J(f \circ g)(\mathbf{x}_0) &= \begin{bmatrix} \frac{\partial}{\partial x_1}(f \circ g)(\mathbf{x}_0) & \frac{\partial}{\partial x_2}(f \circ g)(\mathbf{x}_0) & \cdots & \frac{\partial}{\partial x_3}(f \circ g)(\mathbf{x}_0) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0)) & \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0)) & \cdots & \frac{\partial f}{\partial y_n}(g(\mathbf{x}_0)) \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial g_1}{\partial x_m}(\mathbf{x}_0) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial g_2}{\partial x_m}(\mathbf{x}_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1}(\mathbf{x}_0) & \frac{\partial g_n}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial g_n}{\partial x_m}(\mathbf{x}_0) \end{bmatrix} \end{aligned}$$

de donde podemos ver que, por ejemplo, el j-ésimo elemento de $J(f \circ g)(\mathbf{x}_0)$ (que se obtiene multiplicando la matriz $Jf(g(\mathbf{x}_0))$ por la j-ésima columna de $Jg(\mathbf{x}_0)$) es

$$\begin{aligned}\frac{\partial}{\partial x_j}(f \circ g)(\mathbf{x}_0) &= \frac{\partial f}{\partial y_1}(g(\mathbf{x}_0))\frac{\partial g_1}{\partial x_j}(\mathbf{x}_0) + \frac{\partial f}{\partial y_2}(g(\mathbf{x}_0))\frac{\partial g_2}{\partial x_j}(\mathbf{x}_0) + \cdots + \frac{\partial f}{\partial y_n}(g(\mathbf{x}_0))\frac{\partial g_n}{\partial x_j}(\mathbf{x}_0) \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}(g(\mathbf{x}_0))\frac{\partial g_i}{\partial x_j}(\mathbf{x}_0)\end{aligned}$$

fórmula que ya conocemos de la sección anterior.

Veamos algunos ejemplos.

Ejemplo 3. Consideremos las funciones diferenciables $g\colon \mathbb{R}^2 \to \mathbb{R}^3$ y $f\colon \mathbb{R}^3 \to \mathbb{R}^2$, dadas por

$$\begin{aligned}g(x, y) &= (xy, 5x, y^3)\\ f(x, y, z) &= (3x^2 + y^2 + z^2, 5xyz)\end{aligned}$$

La composición $f \circ g\colon \mathbb{R}^2 \to \mathbb{R}^2$ es diferenciable y su derivada es (donde se denota por $(f \circ g)_1$ y $(f \circ g)_2$ a las funciones coordenadas de $(f \circ g)$).

$$\begin{aligned}J(f \circ g) &= \begin{bmatrix} \frac{\partial}{\partial x}(f \circ g)_1 & \frac{\partial}{\partial y}(f \circ g)_1 \\ \frac{\partial}{\partial x}(f \circ g)_2 & \frac{\partial}{\partial y}(f \circ g)_2 \end{bmatrix} = Jf(g(x, y))Jg(x, y) \\ &= \begin{bmatrix} \frac{\partial f_1}{\partial x}(g(x, y)) & \frac{\partial f_1}{\partial y}(g(x, y)) & \frac{\partial f_1}{\partial z}(g(x, y)) \\ \frac{\partial f_2}{\partial x}(g(x, y)) & \frac{\partial f_2}{\partial y}(g(x, y)) & \frac{\partial f_2}{\partial z}(g(x, y)) \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \end{bmatrix}\end{aligned}$$

Sustituyendo $f_1(x, y, z) = 3x^2 + y^2 + z^2$, $f_2(x, y, z) = 5xyz$, $g_1(x, y) = xy$, $g_2(x, y) = 5x$, $g_3(x, y) = y^3$, nos queda

$$\begin{aligned}&= \begin{bmatrix} 6x & 2y & 2z \\ 5xy & 5xz & 5xy \end{bmatrix}_{g(x,y)} \begin{bmatrix} y & x \\ 5 & 0 \\ 0 & 3y^2 \end{bmatrix} \\ &= \begin{bmatrix} 6(xy) & 2(5x) & 2(y^3) \\ 5(5x)(y^3) & 5(xy)(y^3) & 5(xy)(5x) \end{bmatrix} \begin{bmatrix} y & x \\ 5 & 0 \\ 0 & 3y^2 \end{bmatrix} \\ &= \begin{bmatrix} 6xy^2 + 50x & 6x^2y + 6y^5 \\ 50xy^4 & 100x^2y^3 \end{bmatrix}\end{aligned}$$

Es claro que también podemos llegar a este resultado si antes hacemos explícita la composición

$$\begin{aligned}(f \circ g)(x, y) &= f(g(x, y)) = f(xy, 5x, y^3) \\ &= \big(3(xy)^2 + (5x)^2 + (y^3)^2, 5(xy)(5x)(y^3)\big) \\ &= (3x^2y^2 + 25x^2 + y^6, 25x^2y^4) \qquad \blacksquare\end{aligned}$$

y luego derivamos directamente.

Ejemplo 4. Sea $f: \mathbb{R}^3 \to \mathbb{R}^2$ y $g: \mathbb{R}^3 \to \mathbb{R}^3$ las funciones $f(x, y, z) = (x^2 + 2, x + y^2 + z^3)$, $g(x, y, z) = (x + y + z, xyz, x^2 + y^3)$. Se quiere calcular la derivada de $f \circ g: \mathbb{R}^3 \to \mathbb{R}^2$ en $(1, 1, 1)$. Según la regla de la cadena tenemos

$$J(f \circ g)(1, 1, 1) = Jf(g(1, 1, 1))Jg(1, 1, 1) = Jf(3, 1, 2)Jg(1, 1, 1)$$

Entonces

$$Jf(3, 1, 2) = \begin{bmatrix} 2x & 0 & 0 \\ 1 & 2y & 3x^2 \end{bmatrix}_{(3,1,2)} = \begin{bmatrix} 6 & 0 & 0 \\ 1 & 2 & 12 \end{bmatrix}$$

$$Jg(1, 1, 1) = \begin{bmatrix} 1 & 1 & 1 \\ yz & xz & xy \\ 2x & 3y^2 & 0 \end{bmatrix}_{(1,1,1)} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 3 & 0 \end{bmatrix}$$

de modo que

$$J(f \circ g)(1, 1, 1) = \begin{bmatrix} 6 & 0 & 0 \\ 1 & 2 & 12 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 6 \\ 27 & 39 & 3 \end{bmatrix}$$ ■

Terminamos esta sección estableciendo rigurosamente el teorema que establece la regla de la cadena en el caso general y dando una demostración opcional de él.

Teorema 3.3.1 Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ una función definida en el abierto U de $\mathbb{R}^n$ y $g: V \subseteq \mathbb{R}^m \to \mathbb{R}^n$ una función definida en el abierto V de $\mathbb{R}^m$ tal que $g(V) \subseteq U$. Si g es diferenciable en $\mathbf{x}_0 \in V$ y f es diferenciable en $g(\mathbf{x}_0) \in U$ entonces la función $f \circ g: V \subseteq \mathbb{R}^m \to \mathbb{R}^p$ es diferenciable en $\mathbf{x}_0$ y su derivada viene dada por la matriz

$$J(f \circ g)(\mathbf{x}_0) = Jf(g(\mathbf{x}_0))Jg(\mathbf{x}_0)$$ ■

Demostración. (opcional). Decir que la función $g: V \subseteq \mathbb{R}^m \to \mathbb{R}^n$ es diferenciable en $\mathbf{x}_0$ equivale a decir que la función ρ_1 definida en alguna bola de $\mathbb{R}^m$ con centro en el origen y tomando valores en $\mathbb{R}^n$, definida por

$$g(\mathbf{x}_0 + \mathbf{h}) = g(\mathbf{x}_0) + Jg(\mathbf{x}_0)\mathbf{h} + \|\mathbf{h}\|\rho_1(\mathbf{h})$$

tiene la propiedad de que $\rho_1(\mathbf{0}) = 0$ y $\lim_{\mathbf{h}\to\mathbf{0}} \rho_1(\mathbf{h}) = 0$.

Análogamente, por ser $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ diferenciable en $g(\mathbf{x}_0)$, tenemos

$$f(g(\mathbf{x}_0) + \mathbf{k}) = f(g(\mathbf{x}_0)) + Jf(g(\mathbf{x}_0))\mathbf{k} + \|\mathbf{k}\|\rho_2(\mathbf{k})$$

donde $\rho_2(\mathbf{0}) = 0$ y $\lim_{\mathbf{k}\to\mathbf{0}} \rho_2(\mathbf{k}) = 0$.

Demostraremos que con estas hipótesis la función $f \circ g: V \subseteq \mathbb{R}^m \to \mathbb{R}^p$ es diferenciable en $\mathbf{x}_0$. Tenemos

$$(f \circ g)(\mathbf{x}_0 + \mathbf{h}) = f(g(\mathbf{x}_0 + \mathbf{h})) = f(g(\mathbf{x}_0) + Jg(\mathbf{x}_0)\mathbf{h} + \|\mathbf{h}\|\rho_1(\mathbf{h}))$$

Llamando $\mathbf{k} = Jg(\mathbf{x}_0)\mathbf{h} + \|\mathbf{h}\|\rho_1(\mathbf{h})$ tenemos

$$\begin{aligned}(f \circ g)(\mathbf{x}_0 + \mathbf{h}) &= f(g(\mathbf{x}_0) + \mathbf{k}) = f(g(\mathbf{x}_0)) + Jf(g(\mathbf{x}_0))\mathbf{k} + \|\mathbf{k}\|\rho_2(\mathbf{k}) \\ &= f(g(\mathbf{x}_0)) + Jf(g(\mathbf{x}_0))(Jg(\mathbf{x}_0)\mathbf{h} + \|\mathbf{h}\|\rho_1(\mathbf{h})) + \|\mathbf{k}\|\rho_2(\mathbf{k}) \\ &= (f \circ g)(\mathbf{x}_0) + Jf(g(\mathbf{x}_0))Jg(\mathbf{x}_0)\mathbf{h} + \|\mathbf{h}\|Jf(g(\mathbf{x}_0))\rho_1(\mathbf{h}) + \|\mathbf{k}\|\rho_2(\mathbf{k})\end{aligned}$$

Comparando con la definición de diferenciabilidad de la función $f \circ g$, vemos que todo lo que resta probar es que el residuo

$$r(\mathbf{h}) = \|\mathbf{h}\| Jf(g(\mathbf{x}_0))\rho_1(\mathbf{h}) + \|\mathbf{k}\|\rho_2(\mathbf{k})$$

tiene la propiedad de que $\lim_{\mathbf{h}\to 0} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$. Tenemos

$$\frac{r(\mathbf{h})}{\|\mathbf{h}\|} = Jf(g(\mathbf{x}_0))\rho_1(\mathbf{h}) + \frac{\|\mathbf{k}\|}{\|\mathbf{h}\|}\rho_2(\mathbf{k})$$

Al observar que si $\mathbf{h} \to \mathbf{0}$ entonces $\mathbf{k} \to \mathbf{0}$ y que $\frac{\|\mathbf{k}\|}{\|\mathbf{h}\|}$ es una cantidad que se mantiene acotada en una bola del origen de $\mathbb{R}^m$; vemos que, usando las propiedades de $\rho_1(\mathbf{h})$ y $\rho_2(\mathbf{k})$, se llega a la propiedad deseada de $r(\mathbf{h})$, concluyendo así la demostración del teorema. Q.E.D.

Ejercicios (Capítulo 3, Sección 3)

En cada uno de los ejercicios 1–10 escriba la matriz jacobiana de la función dada en el punto indicado.

1. $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = 256$, $\mathbf{p} = (x_0, y_0)$
2. $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = ax + by + cz$, $\mathbf{p} = (x_0, y_0)$
3. $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y, z) = (a_1x + b_1y, a_2x + b_2y)$, $\mathbf{p} = (x_0, y_0)$
4. $f: \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (x, x + y, x + y + z)$, $\mathbf{p} = (x_0, y_0, z_0)$
5. $f: \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y) = (\operatorname{sen} x, \operatorname{sen} x \cos y, \cos y)$, $\mathbf{p} = (0, \pi/2)$
6. $f: \mathbb{R}^3 \to \mathbb{R}^2$, $f(x, y, z) = \left(\frac{1 + x^2}{1 + z^2}, (z + x^2)(z + y^2)\right)$, $\mathbf{p} = (1, 1, 1)$
7. $f: \mathbb{R}^4 \to \mathbb{R}$, $f(x, y, z, u) = xy^2z^3u^4$, $\mathbf{p} = (1, 1, 1, 1)$
8. $f: \mathbb{R} \to \mathbb{R}^4$, $f(t) = (t, t^2, t^3, t^4)$, $\mathbf{p} = (1)$
9. $f: \mathbb{R}^4 \to \mathbb{R}^4$, $f(x, y, z, u) = (y, u, z, x)$, $\mathbf{p} = (x_0, y_0, z_0, u_0)$
10. $f: \mathbb{R}^2 \to \mathbb{R}^5$, $f(x, y) = (x^y, y^x, e^{xy}, xe^y, ye^x)$, $\mathbf{p} = (1, 2)$
11. Sea $f: \mathbb{R}^n \to \mathbb{R}^m$ una función tal que $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$, $f(c\mathbf{x}) = cf(\mathbf{x})$, para toda $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$, $c \in \mathbb{R}$. Demuestre que f es diferenciable en cualquier punto $\mathbf{p} \in \mathbb{R}^n$. Halle $f'(\mathbf{p})$.
12. Sea $g: \mathbb{R}^2 \to \mathbb{R}^3$ una función diferenciable, que en el punto $\mathbf{p} \in \mathbb{R}^2$ tiene por matriz jacobiana a

$$Jg(\mathbf{p}) = \begin{bmatrix} 2 & -1 \\ 3 & 2 \\ 1 & 7 \end{bmatrix}$$

y sea $f: \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable cuyo gradiente en $g(\mathbf{p}) \in \mathbb{R}^3$ es $\operatorname{grad} g(\mathbf{p}) = (8, 0, -2)$. Demuestre que la función $f \circ g: \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en $\mathbf{p}$. Determine el vector gradiente de esta función en $\mathbf{p}$.

13. Sea $g: \mathbb{R}^2 \to \mathbb{R}^2$ una función diferenciable en el punto $\mathbf{p} \in \mathbb{R}^2$, donde tiene por matriz jacobiana a

$$Jg(\mathbf{p}) = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$$

Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable. Suponga que el gradiente de la función compuesta $f \circ g: \mathbb{R}^2 \to \mathbb{R}$ en el punto $\mathbf{p}$ es $\operatorname{grad} f \circ g(\mathbf{p}) = (1, 1)$. Determine el vector gradiente de f en el punto $g(\mathbf{p})$.

14. Sean $g: \mathbb{R} \to \mathbb{R}^3$, $f: \mathbb{R}^3 \to \mathbb{R}$ dos funciones diferenciables $g(0) = (0, 0, 0)$, $Jg(0) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $Jf(0, 0, 0) = [1\ 1\ 1]$. Demuestre que la función compuesta $f \circ g: \mathbb{R} \to \mathbb{R}$ es creciente en el origen de coordenadas. Determine la ecuación de la recta tangente a la gráfica de esta función en ese punto.

15. Sea $g: \mathbb{R}^3 \to \mathbb{R}^4$ una función diferenciable en el punto $\mathbf{p} \in \mathbb{R}^3$, y $f: \mathbb{R}^4 \to \mathbb{R}^2$ una función diferenciable en el punto $\mathbf{q} = g(\mathbf{p})$. Suponga que

$$Jg(\mathbf{p}) = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 1 & 0 \\ 7 & 0 & 5 \\ 1 & 9 & 8 \end{bmatrix} \qquad Jf(\mathbf{q}) = \begin{bmatrix} -1 & 3 & 2 & 10 \\ 10 & -2 & 2 & -10 \end{bmatrix}$$

Sean $g_1, g_2, g_3, g_4: \mathbb{R}^3 \to \mathbb{R}$ las funciones coordenadas de g, f_1, $f_2: \mathbb{R}^4 \to \mathbb{R}$ las funciones coordenadas de f, y $(f \circ g)_1$, $(f \circ g)_2: \mathbb{R}^3 \to \mathbb{R}$ las funciones coordenadas de la función compuesta $f \circ g: \mathbb{R}^3 \to \mathbb{R}^2$.

a. Escriba los vectores gradientes $\operatorname{grad} g_i(\mathbf{p})$, $i = 1, 2, 3, 4$, y $\operatorname{grad} f_j(\mathbf{q})$, $j = 1, 2$.

b. Determine la ecuación del plano tangente a la superficie de nivel $(f \circ g)_1(x, y, z) = c$ que pasa por $\mathbf{p}$, en este punto.

c. Calcule la derivada direccional de la función $u = (f \circ g)_2(x, y, z)$ en el punto $\mathbf{p}$ en la dirección del vector $\mathbf{v} = (1, 0, 2)$.

En los ejercicios 16–23, sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ una función diferenciable de modo que $f(0, 0) = (0, 0)$. Suponga que la matriz jacobiana de f en $\mathbf{p} = (0, 0)$ es

$$Jf(\mathbf{p}) = \begin{bmatrix} 5 & 2 \\ 1 & 0 \end{bmatrix}$$

Sean f_1, $f_2: \mathbb{R}^2 \to \mathbb{R}$ las funciones coordenadas de f. Determine la matriz jacobiana de las funciones $F: \mathbb{R}^2 \to \mathbb{R}^2$ indicadas en el origen de coordenadas.

16. $F(x, y) = (x, f_1(x, y))$

17. $F(x, y) = (f_2(x, y), y)$

18. $F(x, y) = (x^2 f_1(x, y), y^2 f_2(x, y))$

19. $F(x, y) = f(f(x, y))$

20. $F(x, y) = f(f(f(x, y)))$

21. $F(x, y) = (f_2(x, y), f_1(x, y))$

22. $F(x, y) = (xf_1(x, y) + yf_2(x, y), y f_1(x, y) + xf_2(x, y))$

23. $F(x, y) = (f_1(f_2(x, y), f_2(x, y)), f_2(f_1(x, y), f_1(x, y)))$

En los ejercicios 24–33, sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ una función diferenciable tal que $f(0, 0) = (0, 0)$. Suponga que la matriz jacobiana de f en $\mathbf{p} = (0, 0)$ es

$$Jf(\mathbf{p}) = \begin{bmatrix} 3 & -1 \\ 2 & 2 \end{bmatrix}$$

Sean $f_1, f_2: \mathbb{R}^2 \to \mathbb{R}$ las funciones coordenadas de f. Obtenga las matrices jacobianas de las funciones indicadas en el origen de coordenadas

24. $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = f_1(x, y) + f_2(x, y)$

25. $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = f_1(x, y)f_2(x, y)$

26. $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = \operatorname{sen} f_1(x, y) + \cos f_2(x, y)$

27. $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = \int_{f_1(x,y)}^{f_2(x,y)} g(t)\, dt$, donde $g: \mathbb{R} \to \mathbb{R}$ es una función continua tal que $g(0) = 5$.

28. $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = f_1(x, y) \int_{f_1(x,y)}^{f_2(x,y)} g(t)dt + f_2(x, y) \int_{f_2(x,y)}^{f_1(x,y)} g(t)dt$ donde $g: \mathbb{R} \to \mathbb{R}$ es una función continua tal que $g(0) = 2$.

29. $F: \mathbb{R}^2 \to \mathbb{R}^3$, $F(x, y) = (x, y, f_1(x, y))$

30. $F: \mathbb{R}^2 \to \mathbb{R}^3$, $F(x, y) = (f_1(x, y), f_2(x, y), f_1(x, y) + f_2(x, y))$

31. $F: \mathbb{R}^2 \to \mathbb{R}^3$, $F(x, y) = (f_1(x, f_2(x, y)), f_2(f_1(x, y), y), f_1(x, y)f_2(x, y))$

32. $F: \mathbb{R}^2 \to \mathbb{R}^3$,

$$F(x, y) = \left(3f_1(x, y) + \int_0^{f_2(x,y)} g(t)dt,\ 9f_2(x, y) - 7\int_{f_1(x,y)}^{3} g(t)dt,\ \int_{2f_1(x,y)}^{4f_2(x,y)} g(t)dt\right)$$

donde $g: \mathbb{R} \to \mathbb{R}$ es una función continua tal que $g(0) = 1$.

33. $F: \mathbb{R}^2 \to \mathbb{R}^4$,

$$F(x, y) = \left(xf_1(x, y), yf_2(x, y), x^2(f_1(x, y) + f_2(x, y)), y^2(f_1(x, y) - f_2(x, y))\right)$$

En los ejercicios 34–40, se piede determinar una función diferenciable $g: \mathbb{R}^2 \to \mathbb{R}^2$ cuya matriz jacobiana en el punto $\mathbf{p} \in \mathbb{R}^2$ sea la matriz dada. Se supone dada la función diferenciable $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (f_1(x, y), f_2(x, y))$ que en el punto $\mathbf{p} \in \mathbb{R}^2$ tiene por matriz jacobiana a

$$Jf(\mathbf{p}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

34. $Jg(\mathbf{p}) = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$

35. $Jg(\mathbf{p}) = \begin{bmatrix} 2a & 2b \\ 3c & 3d \end{bmatrix}$

36. $Jg(\mathbf{p}) = \begin{bmatrix} 1 & 0 \\ a & b \end{bmatrix}$

37. $Jg(\mathbf{p}) = \begin{bmatrix} 4 & 5 \\ 8c & 8d \end{bmatrix}$

38. $Jg(\mathbf{p}) = \begin{bmatrix} a+3c & b+3d \\ c-a & d-b \end{bmatrix}$

39. $Jg(\mathbf{p}) = \begin{bmatrix} 4a+5 & 4b-12 \\ a+c+3 & a+c-10 \end{bmatrix}$

40. $Jg(\mathbf{p}) = \begin{bmatrix} f_1(\mathbf{p})a & f_1(\mathbf{p})b \\ f_2(\mathbf{p})c & f_2(\mathbf{p})d \end{bmatrix}$

41. Sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ una función diferenciable que tiene un punto fijo en $\mathbf{p} \in \mathbb{R}^2$, es decir, se tiene $f(\mathbf{p}) = \mathbf{p}$. Sea $A = Jf(\mathbf{p})$ la matriz jacobiana de f en el punto $\mathbf{p}$. Determine una función diferenciable $g: \mathbb{R}^2 \to \mathbb{R}^2$ cuya matriz jacobiana en $\mathbf{p}$ sea $Jg(\mathbf{p}) = A^k$, $k \in \mathbb{N}$.

En cada uno de los ejercicios 42–45, $\phi_1, \psi_1, \phi_2, \psi_2, \phi_3, \psi_3: \mathbb{R} \to \mathbb{R}$ son funciones reales diferenciables definidas en la recta. Determine la matriz jacobiana de las funciones indicadas.

42. a. $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (\phi_1(x), \psi_1(y))$
    b. $g: \mathbb{R}^2 \to \mathbb{R}^2$, $g(x, y) = (\phi_2(x), \psi_2(y))$
    c. $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F = g \circ f$

43. a. $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (\phi_1(x), \psi_1(y) + \psi_2(y))$
    b. $g: \mathbb{R}^2 \to \mathbb{R}^2$, $g(x, y) = (\phi_2(x)\phi_3(x), \psi_3(y))$
    c. $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F = g \circ f$

44. a. $f: \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (\phi_1(x), \phi_2(y), \phi_3(z))$
    b. $g: \mathbb{R}^3 \to \mathbb{R}^3$, $g(x, y, z) = (\psi_1(x), \psi_2(y), \psi_3(z))$
    c. $F: \mathbb{R}^3 \to \mathbb{R}^3$, $F = g \circ f$

45. a. $f: \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (x\phi_1(x), y\phi_2(y), z\phi_3(z))$
    b. $g: \mathbb{R}^3 \to \mathbb{R}^3$, $g(x, y, z) = (x^2\psi_1(x), y^2\psi_2(y), z^2\psi_3(z))$
    c. $F: \mathbb{R}^3 \to \mathbb{R}^3$, $F = g \circ f$

46. Sea $F: \mathbb{R}^n \to \mathbb{R}^n$ la función identidad, $F(\mathbf{x}) = \mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$. Demuestre que la matriz jacobiana de esta función en cualquier punto $\mathbf{p} \in \mathbb{R}^n$ es la matriz identidad.

47. Describa cómo son las funciones diferenciables $F: \mathbb{R}^n \to \mathbb{R}^n$ cuya matriz jacobiana en todo punto $\mathbf{p} \in \mathbb{R}^n$ es una matriz diagonal.

48. Sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ la función $f(x, y) = (ax + by, cx + dy)$, donde $ad - bc = 1$, y sea $g: \mathbb{R}^2 \to \mathbb{R}^2$ la función $g(x, y) = (dx - by, -cx + ay)$. Obtenga la matriz jacobiana de las funciones $f \circ g$, $g \circ f: \mathbb{R}^2 \to \mathbb{R}^2$ en un punto cualquiera $\mathbf{p} \in \mathbb{R}^2$.

49. Sean $g: \mathbb{R}^2 \to \mathbb{R}$, y $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (u(x, y), v(x, y))$ dos funciones diferenciables. Considere la función $g \circ f: \mathbb{R}^2 \to \mathbb{R}$. Demuestre que la derivada de Euler de esta función (ver apéndice II de la sección 12 del capítulo 2) es

$$\mathscr{D}(g \circ f)(x, y) = \frac{\partial g}{\partial x}(f(x, y))\mathscr{D}u(x, y) + \frac{\partial g}{\partial y}(f(x, y))\mathscr{D}v(x, y)$$

o bien, si definimos la derivada de Euler de la función f como $\mathscr{D}f = (\mathscr{D}u, \mathscr{D}v)$, el resultado anterior se puede escribir (en términos del producto punto $\cdot$) como

$$\mathscr{D}(g \circ f)(x, y) = \operatorname{grad} g(f(x, y)) \cdot \mathscr{D}f(x, y)$$

50. Diremos que la función $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (u(x, y), v(x, y))$ es homogénea de grado p si las funciones $u, v: \mathbb{R}^2 \to \mathbb{R}$ lo son. Demuestre que el teorema de Euler sobre funciones homogéneas (establecido para funciones *reales*) se escribe igual para la función f. Es decir, demuestre que si $f: \mathbb{R}^2 \to \mathbb{R}^2$ es homogénea de grado p, entonces $\mathscr{D}f(x, y) = pf(x, y)$.

51. Sean $f, g: \mathbb{R}^2 \to \mathbb{R}^2$ dos funciones diferenciables. Demuestre que la derivada de Euler de la función compuesta $g \circ f: \mathbb{R}^2 \to \mathbb{R}^2$ se escribe como

$$\mathscr{D}(g \circ f)(x, y) = g'(f(x, y))\mathscr{D}f(x, y)$$

en donde el lado derecho de esta expresión se entiende como la multiplicación de la matriz jacobiana (2×2) de g (evaluada en $f(x, y)$) por el vector de $\mathbb{R}^2$ (o la matriz 2×1) $\mathscr{D}f(x, y) = (\mathscr{D}u(x, y), \mathscr{D}v(x, y))$ (en donde $f = (u, v)$).

52. Demuestre que si $f: \mathbb{R}^2 \to \mathbb{R}^2$ es una función homogénea de grado cero y $g: \mathbb{R}^2 \to \mathbb{R}^2$ es una función diferenciable cualquiera, entonces $\mathscr{D}(g \circ f)(x, y) = \mathbf{0}$ (el vector cero de $\mathbb{R}^2$). (Sugerencia: este es un corolario inmediato de los resultados de los dos problemas anteriores).

53. Tome el resultado del ejercicio anterior (junto con el ejemplo 10 de la sección anterior) para concluír que para que la composición $g \circ f$ de dos funciones $g, f: \mathbb{R}^2 \to \mathbb{R}^2$ diferenciables, sea una función homogénea de grado cero es necesario y suficiente que la función f sea homogénea de grado cero. ¿Sigue siendo válido este hecho si la función $f: \mathbb{R}^2 \to \mathbb{R}^2$ es homogénea de grado $p \neq 0$?

3.4 Funciones implícitas (I)

En el capítulo anterior (sección 9) se llamó la atención acerca del hecho de que una curva en el plano, digamos dada como la gráfica de la función $y = f(x)$, se puede ver como una curva de nivel, correspondiente al nivel cero, de una función $z = F(x, y)$. De hecho, esta función debería ser en este caso $F(x, y) = y - f(x)$, de modo que su nivel cero está constituido por los puntos (x, y) tales que $F(x, y) = y - f(x) = 0$, o sea, de modo que $y = f(x)$. Esta misma observación la habíamos hecho con superficies $z = f(x, y)$, las cuales pueden ser vistas como el nivel cero de la función $u = F(x, y, z) = z - f(x, y)$. El punto que llama nuestro interés ante estas consideraciones es el planteamiento recíproco de ellas: por ejemplo, dada la función $z = F(x, y)$, ¿es su nivel cero una curva que se pueda ver como la gráfica de una función $y = f(x)$?, o de otra manera, de la expresión

$F(x, y) = 0$ (que da el nivel cero de F) ¿podemos *despejar* a y en términos de x y dejar así establecida la función $y = f(x)$? Que el nivel cero de $z = F(x, y)$ *sea* una curva en el plano ya ha quedado mostrado en diversas ocasiones que no siempre es cierto (ver, por ejemplo, los ejercicios 11 y 12 de la sección 2 del capítulo 2). Ahora nuestra pregunta es, de hecho, más específica: se quiere saber si el nivel cero de $z = F(x, y)$ es *la gráfica de una función* $y = f(x)$.

Al pensar en la simple función $F(x, y) = x^2 + y^2 - 1$, vemos que la respuesta a la pregunta anterior es *NO*. De la expresión $x^2 + y^2 - 1 = 0$ no podemos establecer *una* función $y = f(x)$ despejando y en términos de x. Es bien sabido que $x^2 + y^2 = 1$ define dos funciones, a saber, $y = f_1(x) = \sqrt{1 - x^2}$ y $y = f_2(x) = -\sqrt{1 - x^2}$.

Consideremos las preguntas anteriores con un poco de más flexibilidad: dada la función $z = F(x, y)$, sea (x_0, y_0) un punto para el cual $F(x_0, y_0) = 0$. De $F(x, y) = 0$, ¿se puede obtener (despejando a y en términos de x) una función $y = f(x)$, *definida en una vecindad de* x_0, tal que $y_0 = f(x_0)$? Obsérvese que lo que se ha logrado con este planteamiento es hacer *LOCAL* la pregunta inicial: dado un punto (x_0, y_0) del nivel cero de la función $z = F(x, y)$ (el cual debe existir para poder empezar la discusión), ¿existe alguna bola de centro (x_0, y_0) en la cual, restringiendo a ella el nivel cero de F, este se pueda ver como la gráfica de una función $y = f(x)$? Cuando tal (bola y tal) función $y = f(x)$ existe, decimos que la función $y = f(x)$ está definida *implícitamente* por la expresión $F(x, y) = 0$, o bien que es una *función implícita* dada en $F(x, y) = 0$. Obsérvese que siendo $y = f(x)$ una función implícita de $F(x, y) = 0$, se debe tener $F(x, f(x)) \equiv 0$ (para toda x en el dominio de f).

Ante esta perspectiva, vemos que las funciones $y = f_1(x) = \sqrt{1 - x^2}$, $y = f_2(x) = -\sqrt{1 - x^2}$ son funciones implícitas definidas en $x^2 + y^2 - 1 = 0$, pues, por ejemplo, el punto $(1/\sqrt{2}, 1/\sqrt{2}) \in \mathbb{R}^2$ es tal que $(1/\sqrt{2})^2 + (1/\sqrt{2})^2 - 1 = 0$ (es decir, pertenece al nivel cero de $F(x, y) = x^2 + y^2 - 1$) y, en una bola de él podemos dibujar la gráfica de la función $y = f(x) = \sqrt{1 - x^2}$

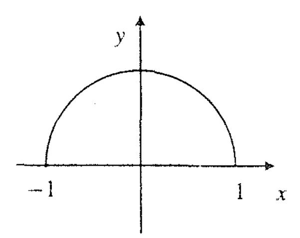


Figura 1. La función $y = f_1(x) = \sqrt{1 - x^2}$.

El problema de la existencia de funciones implícitas $y = f(x)$ dadas por $F(x, y) = 0$ lo resuelve el Teorema de la Función Implícita (TFIm), el cual establece condiciones suficientes (sobre la función $z = F(x, y)$) para que de $F(x, y) = 0$ se pueda obtener alguna función implícita $y = f(x)$. A continuación enunciamos el TFIm en su primera versión (haremos después discusiones similares más generales, y en cada una de ellas aparecerá la correspondiente versión del TFIm).

Teorema 3.4.1 (De la Función Implícita. Primera versión). Considere la función $z = F(x, y)$. Sea $(x_0, y_0) \in \mathbb{R}^2$ un punto tal que $F(x_0, y_0) = 0$. Suponga que la función F tiene derivadas parciales continuas en alguna bola B con centro en (x_0, y_0) y que $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$. Entonces $F(x, y) = 0$ se puede resolver para y en términos de x y definir así una función

$y = f(x)$ con dominio en una vecindad de x_0, tal que $y_0 = f(x_0)$, la cual tiene derivadas continuas en V que pueden calcularse como

$$y' = f'(x) = -\frac{\dfrac{\partial F}{\partial x}(x, y)}{\dfrac{\partial F}{\partial y}(x, y)}, \qquad x \in V$$

■

No es intención de este texto la demostración de este teorema. El lector interesado puede consultarla en alguno de los excelentes libros de Análisis Matemático que hay en el mercado, por ejemplo [EL1] y [CJ]. Quisiéramos en cambio, hacer una serie de observaciones que nos ubiquen en la importancia y contenido del teorema.

En principio debemos tener presente que el TFIm es un *teorema de existencia*: nos habla de que (en ciertas condiciones) *existe* una función $y = f(x)$ definida implícitamente por $F(x, y) = 0$, con tales características. Sin embargo, el teorema nada dice de *cómo* se determina tal función (nos dice que $F(x, y) = 0$ *puede* resolverse para y en términos de x, pero no nos dice cómo hacer el despeje).

EL hecho de que el teorema no nos diga *cómo* establecer explícitamente la función $y = f(x)$, no es simplemente por capricho matemático. La cuestión es que las obstrucciones algebraicas de $F(x, y) = 0$ podrían *no permitir* hacer el despeje de y en términos de x, y... he aquí la importancia del teorema: aun sin poder ("en la práctica") despejar y en términos de x de $F(x, y) = 0$, cumpliendo la función F las hipótesis que el teorema establece, éste *nos garantiza la existencia* de la función $y = f(x)$, y aún más, nos dice cómo calcular su derivada.

Otra observación no menos importante que las anteriores, es que el TFIm es un teorema *local*. Nos asegura la existencia de la función $y = f(x)$, o bien, nos asegura la posibilidad del despeje de y en términos de x a partir de $F(x, y) = 0$, *solamente en las cercanías del punto* (x_0, y_0). Fuera de la vecindad V de que habla el Teorema, éste no se responsabiliza por la existencia de la función f.

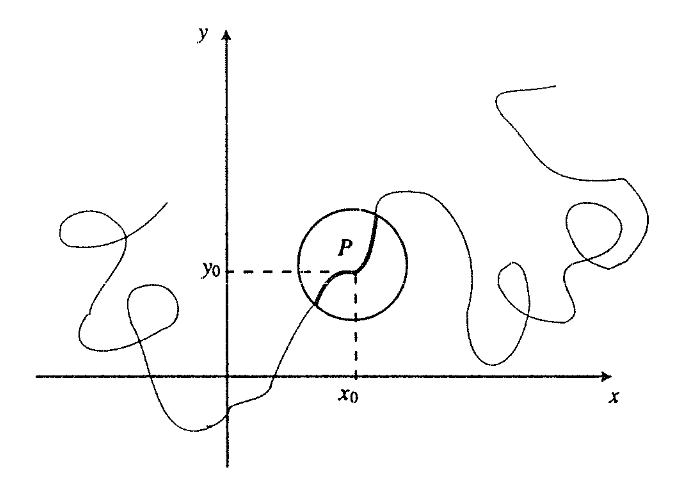


Figura 2. Nivel cero de $z = F(x, y)$.

Observemos, por último, que si asumimos la existencia de la función implícita $y = f(x)$ del TFIm, la fórmula para calcular su derivada se sigue fácilmente de la regla de la cadena. En efecto, para

$x \in V$ se tiene

$$F(x, f(x)) = 0$$

Siendo F y f diferenciables en B y en V respectivamente, podemos aplicar la regla de la cadena, y derivando respecto de x:

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} = 0$$

de donde, como $y = f(x)$ se obtiene fácilmente

$$y' = f'(x) = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}}$$

como lo establece el teorema.

Ejemplo 1. Consideremos la función $F(x, y) = x^2 + y^2 - 1$. Las derivadas parciales de ésta son

$$\frac{\partial F}{\partial x} = 2x \qquad \frac{\partial F}{\partial y} = 2y$$

Estas son continuas siempre. Sea (x_0, y_0) un punto del nivel cero de F, es decir, $x_0^2 + y_0^2 = 1$. Como $\frac{\partial F}{\partial y}(x_0, y_0) = 2y_0$, el TFIm sólo puede aplicarse si $y_0 \neq 0$. Los puntos del nivel cero de F en los que $y_0 = 0$ son $(-1, 0)$ y $(1, 0)$. En estos puntos no podemos concluir del TFIm la existencia de una función $y = f(x)$ en una vecindad de ellos. De hecho, la gráfica de $F(x, y) = 0$ delata el "mal comportamiento" (a la luz de este teorema) de estos puntos.

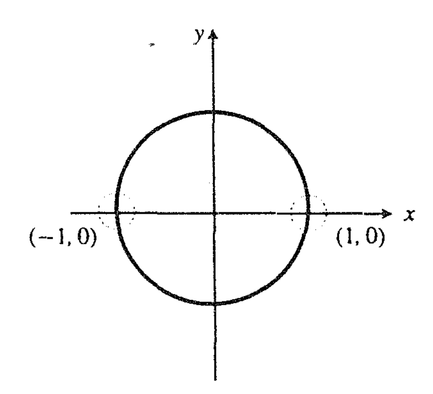


Figura 3. Gráfica del círculo $x^2 + y^2 = 1$.

Para cualquier otro punto (x_0, y_0) tal que $F(x, y) = 0$, con $y_0 \neq 0$, el TFIm garantiza una vecindad V de x_0 en la que se puede definir una función $y = f(x)$ tal que $F(x, f(x)) \equiv 0$, cuya derivada es

$$y' = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = -\frac{2x}{2y} = -\frac{x}{y}, \qquad x \in V$$

Se puede ver que si $y_0 > 0$, tal función es $f(x) = \sqrt{1-x^2}$. De aquí se obtiene

$$f'(x) = \frac{-x}{\sqrt{1-x^2}} = -\frac{x}{y}$$

y si $y_0 < 0$, la función es $f(x) = -\sqrt{1-x^2}$, de donde

$$f'(x) = -\frac{-x}{\sqrt{1-x^2}} = -\frac{x}{-\sqrt{1-x^2}} = -\frac{x}{y}$$

En cualquier caso, la fórmula para la derivada de la función $y = f(x)$ está contenida en la dada por el TFIm. ■

Ejemplo 2. Considere la función $F(x, y) = x^4 - e^{xy^3-1}$. El punto $(1, 1)$ pertenece al nivel cero de F, pues $F(1, 1) = 0$. Las derivadas parciales de F

$$\frac{\partial F}{\partial x} = 4x^3 - y^3e^{xy^3} - 1, \qquad \frac{\partial F}{\partial y} = -3xy^2e^{xy^3} - 1$$

son continuas siempre y además $\frac{\partial F}{\partial y}(1, 1) = -3 \neq 0$. El TFIm nos asegura entonces que en los alrededores de $(1, 1)$, el nivel cero de F se ve como la gráfica de una función $y = f(x)$, y que su derivada es

$$y' = -\frac{4x^3 - y^3e^{xy^3} - 1}{-3xy^2e^{xy^3} - 1} = \frac{4x^3 - y^3e^{xy^3} - 1}{3xy^2e^{xy^3} - 1}$$

Observe que en este caso la función F sí permite hacer el despeje de y en términos de x. De hecho, de $F(x, y) = x^4 - e^{xy^3-1} = 0$, se obtiene $y = f(x) = \left(\frac{1+\ln x^4}{x}\right)^{1/3}$. La derivada y' podría calcularse de esta fórmula explícita y verificar que se llega al mismo resultado obtenido con la fórmula del TFIm. ■

Ejemplo 3. Considere la función $F(x, y) = e^{2y+x} + \operatorname{sen}(x^2 + y) - 1$. En el punto $(0, 0)$ tenemos $F(0, 0) = 0$. Las derivadas parciales de F son

$$\frac{\partial F}{\partial x} = e^{2y+x} + 2x\cos(x^2 + y), \ \frac{\partial F}{\partial y} = 2e^{2y+x} + \cos(x^2 + y)$$

que son siempre continuas. Además $\frac{\partial F}{\partial y}(0, 0) = 3 \neq 0$, de modo que el TFIm garantiza una vecindad V de $x = 0$ en la cual podemos definir una función $y = f(x)$ tal que $F(x, f(x)) = 0$. Obsérvese que en este caso las obstrucciones algebraicas de la función F *no* permiten hacer explícita la función $y = f(x)$, sin embargo tal función *existe* (el TFIm lo dice), y más aún, su derivada es

$$y' = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = -\frac{e^{2y+x} + 2x\cos(x^2 + y)}{2e^{2y+x} + \cos(x^2 + y)}$$ ■

Ejemplo 4. Considere la función $F(x, y) = x^3 + y^3 - 2xy$. De ésta, podríamos hacer explícita una expresión del tipo $y = f(x)$ resolviendo la ecuación cúbica para y. Sin embargo, eso no es

necesario para investigar la curva $F(x, y) = 0$ en los alrededores de un punto concreto, digamos el $(1, 1)$ (ya que $F(1, 1) = 0$). Las derivadas parciales de F son

$$\frac{\partial F}{\partial x} = 3x^2 - 2y, \qquad \frac{\partial F}{\partial y} = 3y^2 - 2x$$

las cuales, son funciones continuas. Se tiene $\frac{\partial F}{\partial y}(1, 1) = 1 \neq 0$. Entonces en los alrededores de $(1, 1)$, la expresión $F(x, y) = 0$ define una función $y = f(x)$, que no necesitamos conocer explícitamente para conocer su derivada

$$y' = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}} = -\frac{3x^2 - 2y}{3y^2 - 2x}$$

que cuando $x = 1$ vale $y'(1) = -1$. Más aún, observamos que $y' = f'(x)$ es una función diferenciable cuando $x = 1$, de modo que podemos calcular y'' derivando respecto de x la expresión de y' recordando que $y = f(x)$. En efecto

$$\begin{aligned} y'' &= \frac{\partial}{\partial x}(y') = \frac{\partial}{\partial x}\left(-\frac{3x^2 - 2y}{3y^2 - 2x}\right) \\ &= -\frac{(3y^2 - 2x)(6x - 2y') - (3x^2 - 2y)(6yy' - 2)}{(3y^2 - 2x)^2} \end{aligned}$$

Cuando $x = 1$, $y = 1$ se tiene $y' = -1$, de modo que

$$y''(1) = -\frac{(3 - 2)(6 - 2(-1)) - (3 - 2)(6(-1) - 2)}{(3 - 2)^2} = -16$$

(si se quisiera la expresión general de y'' en términos de x y y, basta sustituir la expresión de y' previamente encontrada en la de y''). De modo pues que $y'(1) = -1$ y $y''(1) = -16$. Podemos concluir entonces que en los alrededores del punto $(1, 1)$, la curva $F(x, y) = 0$ es decreciente ($y' < 0$) y tiene su gráfica cóncava hacia abajo ($y'' < 0$). La curva completa tiene el siguiente aspecto (llamada Folium de Descartes).

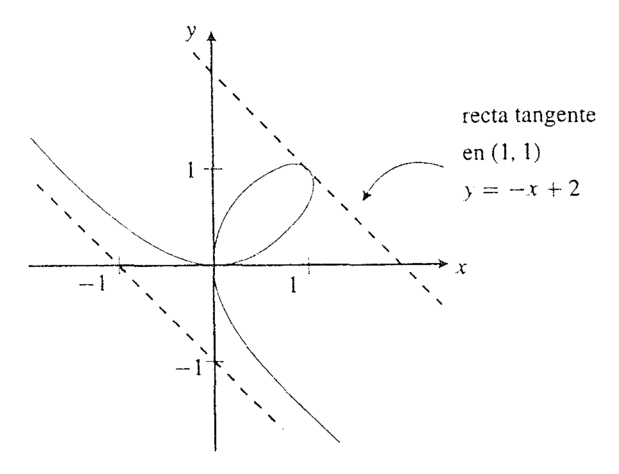


Figura 4. Folium de Descartes. ■

Ejemplo 5. Considere la función $z = F(x, y)$ y el punto $(x_0, y_0) \in \mathbb{R}^2$ tal que $F(x_0, y_0) = 0$. Si F satisface las hipótesis del TFIm, sabemos que en los alrededores de (x_0, y_0), la curva $F(x, y) = 0$ se puede ver como la gráfica de una función $y = f(x)$. ¿Cuál es la ecuación de la recta tangente a la curva $F(x, y) = 0$ en (x_0, y_0)? Todo lo que necesitamos para responder esta pregunta es la pendiente de la recta, y ésta, dada según el TFIm por

$$y'(x_0) = -\frac{\frac{\partial F}{\partial x}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}$$

Entonces la ecuación de la recta tangente procurada es

$$y - y_0 = -\frac{\frac{\partial F}{\partial x}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}(x - x_0)$$

o sea:

$$\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y - y_0) = 0$$

Otro modo de llegar a este resultado es como sigue: siendo la curva en cuestión el nivel cero de $z = F(x, y)$, sabemos que el vector grad $F(x_0, y_0)$ debe ser perpendicular a la recta tangente a la curva en (x_0, y_0). El punto (x, y) estará en tal recta tangente si y sólo si el vector $(x - x_0, y - y_0)$ es perpendicular a grad $F(x_0, y_0) = \left(\frac{\partial F}{\partial x}(x_0, y_0), \frac{\partial F}{\partial y}(x_0, y_0)\right)$, es decir, si su producto punto es cero, obteniendo así la ecuación buscada

$$0 = (x - x_0, y - y_0) \cdot \text{grad } F(x_0, y_0) = \frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y - y_0)$$

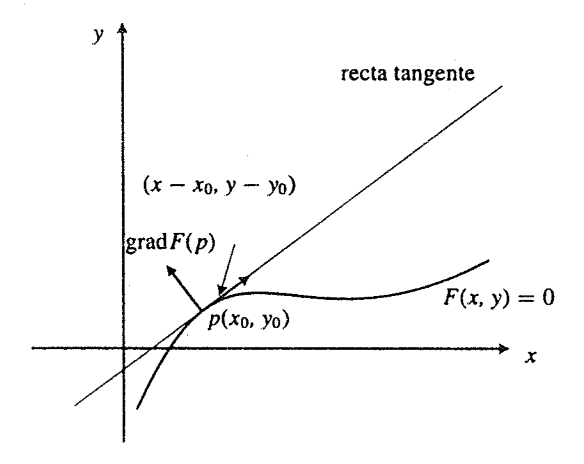


Figura 5. Recta tangente a la curva $F(x, y) = 0$.

El lector puede obtener fácilmente que la ecuación de la recta normal a la curva $F(x, y) = 0$ en (x_0, y_0) es

$$\frac{\partial F}{\partial y}(x_0, y_0)(x - x_0) - \frac{\partial F}{\partial x}(x_0, y_0)(y - y_0) = 0$$ ■

Debemos aclarar que los papeles de las letras x y y en las discusiones anteriores son perfectamente intercambiables. Más bien, suponga que la función $z = F(x, y)$ es tal que en el punto (x_0, y_0) vale cero, que en una bola con centro en (x_0, y_0) tiene derivadas parciales continuas y que su derivada parcial *respecto de x es* es distinta de cero en (x_0, y_0), es decir $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$. El mismo TFIm garantiza la existencia de una función $x = g(y)$ tal que $F(g(y), y) = 0$, para y en una vecindad de y_0. Es decir, garantiza la existencia de una función implícita $x = g(y)$ que pasa por (x_0, y_0), y cuya derivada es

$$\frac{\partial x}{\partial y}(y_0) = g'(y_0) = -\frac{\frac{\partial F}{\partial y}(x_0, y_0)}{\frac{\partial F}{\partial x}(x_0, y_0)}$$

Por ejemplo, para la función analizada en el ejemplo 1, $F(x, y) = x^2 + y^2 - 1$ teníamos que en $(1, 0)$ las derivadas parciales son

$$\frac{\partial F}{\partial x}(1, 0) = 2 \qquad \frac{\partial F}{\partial y}(1, 0) = 0$$

de modo que del TFIm no se concluye la existencia de una función implícita $y = f(x)$, pero *sí* de una función $x = g(y)$ en los alrededores de $(1, 0)$. De hecho esta función es $x = g(y) = \sqrt{1 - y^2}$

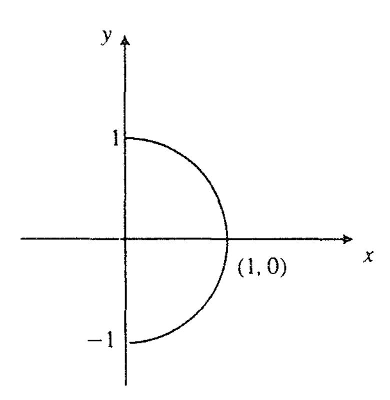


Figura 6. La función implícita $x = \sqrt{1 - y^2}$.

Es claro que si en el punto (x_0, y_0) se tiene

$$\frac{\partial F}{\partial x}(x_0, y_0) \neq 0 \qquad \text{y} \qquad \frac{\partial F}{\partial y}(x_0, y_0) \neq 0$$

el TFIm nos dice que en los alrededores de (x_0, y_0), la gráfica de la curva $F(x, y) = 0$ se puede ver ya sea como la gráfica de una función $y = f(x)$ o bien como la gráfica de una función $x = g(y)$.

Con esta perspectiva, las ecuaciones de las rectas tangente y normal a la curva $F(x, y) = 0$ en un punto (x_0, y_0) de ella, obtenidas en el ejemplo 5, quedan fuera del compromiso de necesitar que

$\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$; según el TFIm basta que $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$ o bien que $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ para que tales ecuaciones hagan sentido.

Este tipo de consideraciones nos hace fijar nuestra atención en puntos (x_0, y_0) de modo que en una bola con centro en ellos, a partir de $F(x, y) = 0$, podemos despejar a x o y en términos de y o x respectivamente y establecer una función con derivada continua $x = g(y)$ ó $y = f(x)$. En tal caso el ejemplo 5 nos proporciona las ecuaciones de las rectas tangente y normal a la curva en (x_0, y_0). Un punto con estas características se llama *punto regular* de la curva $F(x, y) = 0$. Entonces el TFIm nos dice que si $z = F(x, y)$ es una función tal que $F(x_0, y_0) = 0$ y en una bola B con centro en (x_0, y_0) las parciales $\frac{\partial F}{\partial x}$ y $\frac{\partial F}{\partial y}$ son continuas y si $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$ ó $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ (o equivalentemente si $\left(\frac{\partial F}{\partial x}(x_0, y_0)\right)^2 + \left(\frac{\partial F}{\partial y}(x_0, y_0)\right)^2 \neq 0$) entonces (x_0, y_0) es un punto regular de la curva $F(x, y) = 0$.

Ejemplo 6. Considere la función $F(x, y) = 9x^2 + 4y^2 - 18x - 8y - 23$ que tiene derivadas parciales

$$\frac{\partial F}{\partial x} = 18x - 18 \qquad \frac{\partial F}{\partial y} = 8y - 8$$

las cuales son continuas. Si (x_0, y_0) es tal que $F(x_0, y_0) = 0$, entonces es claro que $\left(\frac{\partial F}{\partial x}(x_0, y_0)\right)^2 + \left(\frac{\partial F}{\partial y}(x_0, y_0)\right)^2 \neq 0$. Es decir, todos los puntos de la curva $F(x, y) = 0$ son regulares: alrededor de cualquier punto de la curva se puede ver como la gráfica de una función $y = f(x)$ o $x = g(y)$. De hecho $F(x, y) = 0$ es una elipse con semiejes 2 y 3 y centro en $(1, 1)$

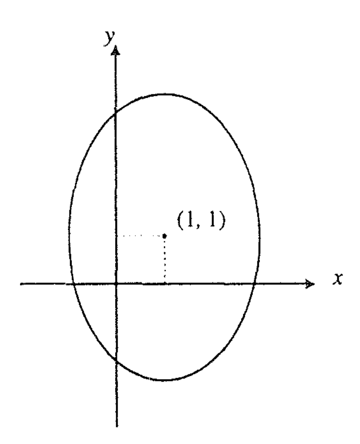


Figura 7. Elipse del ejemplo 6. ■

Si el punto (x_0, y_0) no es un punto regular de $F(x, y) = 0$, se dice que es un *punto singular*. Pasemos ahora a discutir una generalización natural de los asuntos estudiados en esta sección. Si tomamos ahora una función $F: \mathbb{R}^{n+1} \to \mathbb{R}$, digamos $z = F(x_1, x_2, \ldots, x_n, y)$ y consideremos el nivel cero de ella $F(x_1, x_2, \ldots, x_n, y) = 0$, ahora la pregunta es cuándo este nivel se puede ver como la gráfica de una función $y = f(x_1, x_2, \ldots, x_n)$ (vea definición de gráfica de una función de n variables en la sección 2 del capítulo anterior). Es decir, cuándo, de la expresión $F(x_1, x_2, \ldots, x_n, y) = 0$ se puede despejar "y" en términos de las n variables $x_1, x_2, \ldots, x_n$, y formar así una función $y = f(x_1, x_2, \ldots, x_n)$. El caso que ya se discutió exhaustivamente es $n = 1$. La discusión en esta nueva situación con n arbitraria, se copia casi textualmente del caso $n = 1$.

Sea $p = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{y}) \in \mathbb{R}^{n+1}$ en un punto para el cual $F(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, y) = 0$. Si existe una bola B (en $\mathbb{R}^{n+1}$) con centro en **p**, tal que restringiendo a $F(x_1, x_2, \ldots, x_n, y) = 0$ a B, ésta se ve como la gráfica de una función $y = f(x_1, x_2, \ldots, x_n)$ (con $\bar{y} = f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$), decimos que esta función f es una *función implícita* dada en $F(x_1, x_2, \ldots, x_n, y) = 0$.

El siguiente TFIm nos dice en qué condiciones se puede esperar la existencia de funciones implícitas.

Teorema 3.4.2 (De la Función Implícita. Segunda versión). Considere la función $z = F(x_1, x_2, \ldots, x_n, y)$. Sea $\mathbf{p} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{y}) \in \mathbb{R}^{n+1}$ un punto tal que $F(\mathbf{p}) = 0$. Suponga que la función F tiene derivadas parciales $\frac{\partial F}{\partial x_i}$, $i = 1, 2, \ldots, n$ y $\frac{\partial F}{\partial y}$ continuas en alguna bola B con centro en **p** y que $\frac{\partial F}{\partial y}(\mathbf{p}) \neq 0$. Entonces $F(x_1, x_2, \ldots, x_n, y) = 0$ puede resolverse para y en términos de x y definir así una vecindad V (de $\mathbb{R}^n$) del punto $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$, una función $y = f(x_1, x_2, \ldots, x_n)$ la cual tiene derivadas parciales continuas en V que se pueden calcular con las fórmulas

$$\frac{\partial f}{\partial x_i}(x_1, x_2, \ldots, x_n) = -\frac{\frac{\partial F}{\partial x_i}(x_1, x_2, \ldots, x_n, y)}{\frac{\partial F}{\partial y}(x_1, x_2, \ldots, x_n, y)} \qquad (x_1, x_2, \ldots, x_n) \in V$$ ■

Veamos algunos ejemplos.

Ejemplo 7. Sea la función $F(x, y, z) = x^2 + y^2 + z^2 - 3$. El punto $\mathbf{p} = (1, 1, 1) \in \mathbb{R}^3$ es tal que $F(\mathbf{p}) = 0$. Las derivadas parciales de F son

$$\frac{\partial F}{\partial x} = 2x, \qquad \frac{\partial F}{\partial y} = 2y, \qquad \frac{\partial F}{\partial z} = 2z$$

Estas siempre son continuas. En el punto **p**, se tiene $\frac{\partial F}{\partial z}(\mathbf{p}) = 2 \neq 0$. El TFIm dice entonces que en los alrededores del punto **p**, $F(x, y, z) = 0$ puede verse como la gráfica de una función $z = f(x, y)$ que tiene por derivadas parciales a

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{2x}{2z} = -\frac{x}{z}$$

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{2y}{2z} = -\frac{y}{z}$$

De hecho, es claro que tal función f es $f(x, y) = \sqrt{3 - x^2 - y^2}$, geométricamente este ejemplo es equivalente al ejemplo 1, sólo que con una dimensión más: $F(x, y, z) = 0$ representa una esfera con centro en el origen y radio $\sqrt{3}$, la cual globalmente no es la gráfica de función $z = f(x, y)$ alguna. Pero alrededor del punto $(1, 1, 1)$ de tal esfera, esto se puede ver como la gráfica de la función $f(x, y) = \sqrt{3 - x^2 - y^2}$. ■

Ejemplo 8. Sea $F(x, y, z) = x + y + z - ze^z$. Las derivadas parciales de esta función son

$$\frac{\partial F}{\partial x} = 1, \quad \frac{\partial F}{\partial y} = 1, \quad \frac{\partial F}{\partial z} = 1 - e^z(z + 1)$$

Si el punto $\mathbf{p} = (x_0, y_0, z_0) \in \mathbb{R}^3$ es tal que $x_0 + y_0 + z_0 - z_0e^{z_0} = 0$ y $z \neq 0$, entonces, puesto que $\frac{\partial F}{\partial z}(\mathbf{p}) \neq 0$, el TFIm sugiere que podamos despejar z en términos de x y y y establecer así una función $z = f(x, y)$ (con $z_0 = f(x_0, y_0)$) de modo que su gráfica en los alrededores de $\mathbf{p}$ coincide con $F(x, y, z) = 0$. Las parciales de la función f son

$$\frac{\partial f}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{1}{1 - e^z(z + 1)}, \quad \frac{\partial f}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{1}{1 - e^z(z + 1)}$$

■

Ejemplo 9. Sea $F(x, y, z, u, v) = x^2 + 2y^2 + 3z^2 + 4u^2v + e^{u+v} - 1$. Se tiene $F(0, 0, 0, 0, 0) = 0$. Las parciales de F son

$$\frac{\partial F}{\partial x} = 2x \quad \frac{\partial F}{\partial y} = 4y \quad \frac{\partial F}{\partial z} = 6z \quad \frac{\partial F}{\partial u} = 8uv + e^{u+v} \quad \frac{\partial F}{\partial v} = 4u^2 + e^{u+v}$$

que son funciones continuas. En el origen (de $\mathbb{R}^5$) la parcial $\frac{\partial F}{\partial v}$ vale 1 ($\neq 0$). EL TFIm asegura entonces que alrededor del origen $F(x, y, z, u, v) = 0$ se puede ver como la gráfica de una función $v = f(z, y, z, u)$ (que no es posible hacer explícita en este caso) cuyas derivadas en el origen son

$$\frac{\partial v}{\partial x}(0, 0, 0, 0, 0) = -\frac{\frac{\partial F}{\partial x}(0, 0, 0, 0, 0)}{\frac{\partial F}{\partial v}(0, 0, 0, 0, 0)} = -\frac{0}{1} = 0$$

Análogamente

$$\frac{\partial v}{\partial y}(0, 0, 0, 0, 0) = 0, \quad \frac{\partial v}{\partial z}(0, 0, 0, 0, 0) = 0, \quad \frac{\partial v}{\partial u}(0, 0, 0, 0, 0) = -1$$

■

Ejemplo 10. En el cálculo muchas veces se suele ser poco riguroso con las funciones implícitas. Es común encontrarse con expresiones del tipo "... la función $y = f(x)$ dada implícitamente por $F(x, y) = 0$ ..." Sin embargo, habiendo discutido con el debido detalle lo que aquí se ha presentado sobre funciones implícitas, podemos, con un poco de buena voluntad, entender lo que está detrás de esta afirmación y que nos dice: "como la función $z = F(x, y)$ satisface en el punto (x_0, y_0) las hipótesis del TFIm, es decir, **1.** $F(x_0, y_0) = 0$; **2.** en una bola con centro en (x_0, y_0) las derivadas parciales $\frac{\partial F}{\partial x}$ y $\frac{\partial F}{\partial y}$ son continuas, y **3.** $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$, entonces considere la función implícita $y = f(x)$ que el TFIm asegura que existe ...". Es claro que esto es muy largo para escribirlo cada vez que nos referimos a alguna función implícita dada por $F(x, y) = 0$. Es por eso que, en lo que sigue, también nos daremos las libertades pertinentes que nos simplifiquen la comunicación. Este (único) preámbulo, lo hacemos para que haga sentido el enunciado de este y muchos otros ejercicios y discusiones que aparecerán más adelante en este y otros capítulos.

Considere, pues, la función $z = f(x, y)$ dada implícitamente por $F(x, y, z) = xyz - e^z = 0$. Verifiquemos que ésta satisfaga la conclusión del Teorema de Schwarz sobre la igualdad de las derivadas parciales cruzadas en el punto $\mathbf{p} = (e^2, \frac{1}{2}, 2)$

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{yz}{xy - e^z}, \quad \frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{xz}{xy - e^z}$$

que en el punto $\mathbf{p}$ son

$$\frac{\partial z}{\partial x} = -\frac{(\frac{1}{2})(2)}{\frac{1}{2}e^2 - e^2} = 2e^{-2}, \quad \frac{\partial z}{\partial y} = -\frac{2e^2}{\frac{1}{2}e^2 - e^2} = 4$$

Al calcular las derivadas de segundo orden, debemos tener siempre presente que $z = f(x, y)$ y aplicar entonces la regla de la cadena

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}\left(-\frac{yz}{xy - e^z}\right)$$
$$= -\frac{(xy - e^z)\left(y\frac{\partial z}{\partial y} + z\right) - yz\left(x - e^z\frac{\partial z}{\partial y}\right)}{(xy - e^z)^2}$$

que en el punto $\mathbf{p}$ es

$$\frac{\partial^2 z}{\partial y \partial x}(\mathbf{p}) = -4e^{-2}$$

Por otro lado

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}\left(-\frac{xz}{xy - e^z}\right) = -\frac{(xy - e^z)z - xyz}{(xy - e^z)^2} - \frac{(xy - e^z)y + yze^z}{(xy - e^z)^2}\frac{\partial z}{\partial x}$$
$$= -\frac{(xy - e^z)\left(x\frac{\partial z}{\partial x} + z\right) - xz\left(y - e^z\frac{\partial z}{\partial x}\right)}{(xy - e^z)^2}$$

que en el punto $\mathbf{p}$ es

$$\frac{\partial^2 z}{\partial x \partial y}(\mathbf{p}) = -4e^{-2}$$

y así $\frac{\partial^2 z}{\partial y \partial x}(\mathbf{p}) = \frac{\partial^2 z}{\partial x \partial y}(\mathbf{p})$, como queríamos comprobar. ■

Nuevamente debemos advertir que siendo $u = F(x, y, z)$ una función que cumple las hipótesis del TFIm, las letras x, y, z juegan papeles completamente intercambiables al considerar la expresión $F(x, y, z) = 0$. Es decir, de $F(x, y, z) = 0$ podremos despejar una de las variables en términos de las otras dos restantes siempre que la derivada parcial de F respecto de esa variable sea distinta de cero (localmente). Entonces, si $\mathbf{p} = (x_0, y_0, z_0)$ es un punto para el cual alguna de las derivadas parciales $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ o $\frac{\partial F}{\partial z}$ es no nula, el TFIm nos dice que en los alrededores de $\mathbf{p}$ podemos ver a $F(x, y, z) = 0$ como la gráfica de una función $x = h(y, z)$, $y = g(x, z)$ o $z = f(x, y)$, respectivamente.

Por ejemplo, considere la expresión $F(x, y, z) = 0$ donde $F(x_0, y_0, z_0) = 0$, F tiene derivadas parciales continuas en una bola alrededor de $\mathbf{p} = (x_0, y_0, z_0)$ y alguna de ellas no se anula en $\mathbf{p}$, digamos que $\frac{\partial F}{\partial z}(\mathbf{p}) \neq 0$. El TFIm nos dice entonces que en los alrededores de $\mathbf{p}$ podemos ver a $F(x, y, z) = 0$ como la gráfica de una función $z = f(x, y)$. ¿Cuál es la ecuación del plano tangente a esta gráfica en $\mathbf{p}$? Según lo discutido en la sección 10 del capítulo anterior, todo lo que necesitamos determinar son las derivadas $\frac{\partial f}{\partial x}(x_0, y_0)$ y $\frac{\partial f}{\partial y}(x_0, y_0)$, y éstas las calculamos con la ayuda del TFIm que nos dice que

$$\frac{\partial f}{\partial x}(x_0, y_0) = -\frac{\frac{\partial F}{\partial x}(\mathbf{p})}{\frac{\partial F}{\partial z}(\mathbf{p})} \qquad \frac{\partial f}{\partial y}(x_0, y_0) = -\frac{\frac{\partial F}{\partial y}(\mathbf{p})}{\frac{\partial F}{\partial z}(\mathbf{p})}$$

Así, la ecuación requerida del plano tangente es

$$z - z_0 = -\frac{\frac{\partial F}{\partial x}(\mathbf{p})}{\frac{\partial F}{\partial z}(\mathbf{p})}(x - x_0) - \frac{\frac{\partial F}{\partial y}(\mathbf{p})}{\frac{\partial F}{\partial z}(\mathbf{p})}(y - y_0)$$

o sea

$$\frac{\partial F}{\partial x}(\mathbf{p})(x - x_0) + \frac{\partial F}{\partial y}(\mathbf{p})(y - y_0) + \frac{\partial F}{\partial z}(\mathbf{p})(z - z_0) = 0$$

Pudimos haber llegado a establecer esta ecuación haciendo uso de que siendo $F(x, y, z) = 0$ una superficie de nivel de la función F, el vector grad $F(\mathbf{p}) = \left(\frac{\partial F}{\partial x}(\mathbf{p}), \frac{\partial F}{\partial y}(\mathbf{p}), \frac{\partial F}{\partial z}(\mathbf{p})\right)$ es el vector ortogonal a dicha superficie en $\mathbf{p}$, y entonces, el vector $(x - x_0, y - y_0, z - z_0)$ es ortogonal a grad $F(\mathbf{p})$, o sea, si y sólo si el producto de estos dos vectores es cero.

$$\begin{aligned} 0 &= (x - x_0, y - y_0, z - z_0) \cdot \left(\frac{\partial F}{\partial x}(\mathbf{p}), \frac{\partial F}{\partial y}(\mathbf{p}), \frac{\partial F}{\partial z}(\mathbf{p})\right) \\ &= \frac{\partial F}{\partial x}(\mathbf{p})(x - x_0) + \frac{\partial F}{\partial y}(\mathbf{p})(y - y_0) + \frac{\partial F}{\partial z}(\mathbf{p})(z - z_0) \end{aligned}$$

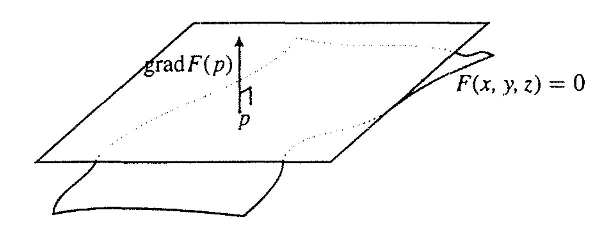


Figura 8. Plano tangente de la superficie $F(x, y, z) = 0$.

Observamos también que la ecuación del plano tangente anteriormente planteada hace perfecto sentido aun cuando $\frac{\partial F}{\partial z}(\mathbf{p}) = 0$. Lo que se requiere es que no se anulen simultáneamente en $\mathbf{p}$ las tres derivadas parciales de la función F, pues, como dijimos, en tal caso el TFIm nos dice que en los alrededores de $\mathbf{p}$ la superficie $F(x, y, z) = 0$ se puede ver como la gráfica de una función (diferenciable) de dos variables, y por lo tanto se le puede asociar un plano tangente, cuya ecuación es la establecida anteriormente. (Ver ejercicios 6 y 7 de la sección 10 del capítulo 2).

Ejemplo 11. Considere la superficie en $\mathbb{R}^3$ definida implícitamente por

$$F(x, y, z) = xyz + \ln(xyz) - z = 0$$

Hallar la ecuación del plano tangente en $\mathbf{p} = (1, 1, 1)$. Se tiene

$$\frac{\partial F}{\partial x} = yz + \frac{1}{x}, \quad \frac{\partial F}{\partial y} = xz + \frac{1}{y}, \quad \frac{\partial F}{\partial z} = xy + \frac{1}{z} - 1$$

de modo que evaluando en $\mathbf{p}$

$$\frac{\partial F}{\partial x}(\mathbf{p}) = 2, \quad \frac{\partial F}{\partial y}(\mathbf{p}) = 2, \quad \frac{\partial F}{\partial z}(\mathbf{p}) = 1$$

y así, la ecuación del plano tangente procurada es

$$2(x - 1) + 2(y - 1) + (z - 1) = 0$$

o sea

$$2x + 2y + z = 5$$

■

Ejemplo 12. Hallar la ecuación del plano tangente a la superficie dada implícitamente por

$$F(x, y, z) = 36x^2 + 9y^2 + 4z^2 - 72x - 36y - 24z + 72 = 0$$

en el punto $\mathbf{p} = (1, 4, 3)$. Se tiene

$$\frac{\partial F}{\partial x} = 72x - 72, \quad \frac{\partial F}{\partial y} = 18y - 36, \quad \frac{\partial F}{\partial z} = 8z - 24$$

de modo que en el punto $\mathbf{p}$ obtenemos

$$\frac{\partial F}{\partial x}(\mathbf{p}) = 0, \quad \frac{\partial F}{\partial y}(\mathbf{p}) = 36, \quad \frac{\partial F}{\partial y}(\mathbf{p}) = 0$$

(Obsérvese que en este caso podemos aplicar el TFIm para concluir que la superficie $F(x, y, z) = 0$ se ve en los alrededores del punto $\mathbf{p}$ como la gráfica de una función del tipo $y = g(x, z)$). El plano tangente procurado tiene por ecuación a

$$0(x - 1) + 4(y - 4) + 0(z - 3) = 0$$

o sea

$$y = 4$$

■

Ejercicios (Capítulo 3, Sección 4)

En los ejercicios 1–5 se dan funciones $F(x, y)$. Verifique en cada caso que F satisface las hipótesis del teorema de la función implícita (en algún punto $\mathbf{p}$ del nivel cero de F), y obtenga la derivada de la función $y = f(x)$ definida por el nivel cero de F. En cada caso es posible hacer explícita esta última función. Hágalo y obtenga de nuevo y' derivando directamente la función $y = f(x)$ despejada.

1. $F(x, y) = 8x + 10y - 2$

2. $F(x, y) = 2xy + y - 4$

3. $F(x, y) = x^2 + 3x^3 + 8xy^3 - 20$

4. $F(x, y) = x - 2 - 5e^x + e^y$

5. $F(x, y) = 2x^2 + 4x - 9\ln(1 + 4x^2 + 3y^3)$

En los ejercicios 6–10, se da el nivel cero de una cierta función $F(x, y)$. Compruebe que esta función satisface las hipótesis del teorema de la función implícita en el punto indicado (perteneciente al nivel cero de F). Obtenga la derivada de la función $y = f(x)$ en el punto dado.

6. $F(x, y) = x^2y + 3x^2 - 2y^2 - 2y = 0$, $\mathbf{p} = (1, 1)$

7. $F(x, y) = \operatorname{sen} x + \cos y + 2y - \pi = 0$, $\mathbf{p} = (0, \pi/2)$

8. $F(x, y) = y\ln(x^2 + y^2) - 2xy = 0$, $\mathbf{p} = (0, 1)$

9. $F(x, y) = xe^x + ye^y - 2x - 2y = 0$, $\mathbf{p} = (0, 0)$

10. $F(x, y) = x^y + y^x - 2xy = 0$, $\mathbf{p} = (2, 2)$

En los ejercicios 11–15 se da el nivel cero de una cierta función $F(x_1, x_2, y)$. Compruebe que éste define implícitamente una función $y = f(x_1, x_2)$ en una vecindad del punto $\mathbf{p}$ dado perteneciente al nivel cero de F. Obtenga las derivadas parciales de la función f en $\mathbf{p}$.

11. $F(x_1, x_2, y) = x_1 \operatorname{sen}^2 x_2 + x_1 - 3x_2 + y = 0$, $\mathbf{p} = (0, 0, 0)$

12. $F(x_1, x_2, y) = x_1 \ln(1 + x_2) + ye^{4y} = 0$, $\mathbf{p} = (0, 0, 0)$

13. $F(x_1, x_2, y) = y \arctan(1 - y^2) + 3x_1 + 5y - 8x_2^3 = 0$, $\mathbf{p} = (1, 1, 1)$

14. $F(x_1, x_2, y) = x_1(x_2 + e^y) + 5y - 2 = 0$, $\mathbf{p} = (1, 1, 0)$

15. $F(x_1, x_2, y) = x_1x_2ye^y \ln y - 3x_1 + 3x_2 = 0$, $\mathbf{p} = (1, 1, 1)$

16. Demuestre que cualquier nivel constante de una función lineal $F: \mathbb{R}^n \to \mathbb{R}$, $F(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$, donde $a_1a_2 \ldots a_n \neq 0$, define siempre funciones implícitas $f_i: \mathbb{R}^{n-1} \to \mathbb{R}$, $x_i = f_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, las cuales también son funciones lineales. Determine las derivadas parciales de las funciones f_i. ¿Qué importancia tiene la suposición de que el producto $a_1a_2 \cdots a_n$ sea distinto de cero en el resultado establecido en este ejercicio?

17. Considere la función $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = ax^2 + 2bxy + cy^2 + d$, en donde $d \neq 0$. ¿En qué condiciones es posible trazar una recta tangente a la gráfica de $F(x, y) = 0$ en cualquier punto de ella?

18. Sea $y = f(x)$ una función dos veces diferenciable definida implícitamente por $F(x, y) = 0$. Demuestre que la segunda derivada y"(x) viene dada por

$$y''(x) = -\frac{\dfrac{\partial^2 F}{\partial x^2}\left(\dfrac{\partial}{\partial y}\right)^2 - 2\dfrac{\partial^2 F}{\partial x \partial y}\dfrac{\partial F}{\partial x}\dfrac{\partial F}{\partial y} + \dfrac{\partial^2 F}{\partial y^2}\left(\dfrac{\partial F}{\partial x}\right)^2}{\left(\dfrac{\partial F}{\partial y}\right)^3}$$

(todas las derivadas parciales de F calculadas en $(x, f(x))$).

19. Suponga que la expresión $F(x, y) = F(y, x)$, en que $F: \mathbb{R}^2 \to \mathbb{R}$ es una función diferenciable, define implícitamente una función diferenciable $y = f(x)$. Derivando respecto de x, nos queda que

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}y' = \frac{\partial F}{\partial x}y' + \frac{\partial F}{\partial y}$$

de donde $y' = 1$, lo cual es, en general, falso. Encuentre el error en este razonamiento. Determine la expresión correcta para la derivada $f'(x)$.

20. Suponga que la expresión $F(x, \phi(x), y) = 0$, donde $F: \mathbb{R}^3 \to \mathbb{R}$, y $\phi: \mathbb{R} \to \mathbb{R}$ son funciones diferenciables, define implícitamente una función diferenciable $y = f(x)$. Halle $f'(x)$. (Sugerencia: derive respecto de x la expresión $F(x, \phi(x), y) = 0$, usando adecuadamente la regla de la cadena).

21. Suponga que la expresión $F(\phi(x), \psi(x), y^3) = 0$, donde $F: \mathbb{R}^3 \to \mathbb{R}$, $\phi, \psi: \mathbb{R} \to \mathbb{R}$ son funciones de clase $\mathscr{C}^1$, define implícitamente una función diferenciable $y = f(x)$. Halle $f'(x)$.

22. Sean $G, F: \mathbb{R}^2 \to \mathbb{R}$ funciones diferenciables. Suponga que la expresión $G(F(x, y), F(y, x)) = 0$ define implícitamente una función diferenciable $y = f(x)$. Determine la expresión para $f'(x)$.

En los ejercicios 23–26 se da el nivel cero de una función diferenciable $F: \mathbb{R}^4 \to \mathbb{R}$, $w = F(x, y, z, u)$, y un punto **p** perteneciente a este nivel. Diga en cada caso si en los alrededores del punto **p** es posible ver la gráfica de F como la gráfica de una función diferenciable del tipo: **a.** $u = u(x, y, z)$, **b.** $z = z(x, y, u)$, **c.** $y = y(x, z, u)$, y/o, **d.** $x = x(y, z, u)$. En cada caso (en el que tal función exista), determine sus derivadas parciales en el punto **p**.

23. $x^2 + y^2 + z^2 + u^2 - 4 = 0$, $\mathbf{p} = (1, 1, 1, 1)$

24. $xyzu + x^3 - 5yz^2 + 8u - 8z = 0$, $\mathbf{p} = (0, 0, 1, 1)$

25. $\ln(1 + x^2 + y^2) + 3z - 8u = 0$, $\mathbf{p} = (0, 0, 0, 0)$

26. $x \operatorname{sen} x + y \operatorname{sen} y + z \operatorname{sen} z + u \operatorname{sen} u = 0$, $\mathbf{p} = (0, 0, 0, 0)$

27. Determine la derivada direccional de la función $z = f(x, y)$ definida implícitamente por $x \tan y - ze^z = 0$ en el punto $\mathbf{p} = (0, \pi/4, 0)$ en la dirección del vector $u = (2, 1)$.

28. Determine la derivada direccional de la función $u = f(x, y, z)$ definida implícitamente por $u + ye^u + x + 3z = 0$ en el origen de coordenadas en la dirección del vector $\mathbf{v} = (1, -1, -1)$.

29. Hallar la dirección de mayor crecimiento de la función $z = f(x, y)$ dada implícitamente por $\arctan(x + y + z) + 3xyz + z = 0$ en el origen de coordenadas.

30. Considere la superficie $-2x^2 + 64x - 4y^2 + 64y + z^2 - 768 = 0$. ¿En qué punto de ella no es posible trazar un plano tangente? Explique.

31. Repita el ejercicio anterior con la superficie

$$-x^2 + x(2z + 10) - y^2 + y(2z + 14) + z^2 + 8z + 6 = 0$$

En los ejercicios 32–34 considere la función $z = f(x, y)$ definida implícitamente por la expresión dada $F(x, y, z) = 0$. Calcule las derivadas parciales de segundo orden de la función f.

32. $x^2y - 3z + 8yz^3 = 0$

33. $\operatorname{sen}(xy) + z + \operatorname{sen} z = 0$

34. $xe^x + ye^y + ze^z - 3e = 0$, en el punto $(1, 1, 1)$.

35. Suponga que la expresión $F(x, y) = 0$ determina funciones diferenciables $x = f(y)$, $y = g(x)$. Demuestre que $f'(y)g'(x) = 1$. ¿Qué relación tiene este hecho con el teorema de la función inversa del curso de cálculo de funciones de una variable?

36. Suponga que la expresión $F(x, y, z) = 0$ determina implícitamente funciones diferenciables $x = x(y, z)$, $y = y(x, z)$, $z = z(x, y)$. Demuestre que

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1$$

En los ejercicios 37–42, $F: \mathbb{R}^3 \to \mathbb{R}$ es una función de clase $\mathscr{C}^1$. Suponga que la expresión dada determina implícitamente la función de clase $\mathscr{C}^1$ $z = f(x, y)$. Hallar las derivadas parciales de esta función.

37. $F(y, x, z) = 0$

38. $F(z, y, x) = 0$

39. $F(x, y, z) + F(x, z, y) + F(z, x, y) = 0$

40. $F(a_1x + b_1y + c_1z, a_2x + b_2y + c_2z, a_3x + b_3y + c_3z) = 0$

41. $xF(y, y, z) + yF(x, x, z) + zF(x, x, x) = 0$

42. $F(x \operatorname{sen} y \cos z, y \operatorname{sen} z \cos y, z \operatorname{sen} x \cos y) = 0$

43. $F\left(\frac{x}{y}, \frac{y}{z}, \frac{z}{x}\right) = 0$

44. Sea $F: \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathscr{C}^2$. Suponga que la expresión $F(x + z, x) = 0$ define implícitamente una función $z = f(x, y)$ de clase $\mathscr{C}^2$. Determine las derivadas parciales de segundo orden de esta función.

45. Suponga que la expresión

$$\int_{xz}^{y+z} g(t)\,dt + \int_{3x+y}^{z^2} h(t)\,dt = 0$$

donde $g, h: \mathbb{R} \to \mathbb{R}$ son funciones continuas, define implícitamente una función diferenciable $z = f(x, y)$. Halle sus derivadas parciales.

3.5 Funciones implícitas (II)

Vamos a estudiar ahora una generalización de los temas analizados en la sección anterior. Comencemos por considerar una situación elemental que nos introduzca en el tipo de ideas que manejaremos en esta sección. Tomemos el siguiente sistema lineal de 2 ecuaciones con las variables u, v, x, y

$$\begin{aligned} au + bv - k_1x &= 0 \\ cu + dv - k_2y &= 0 \end{aligned}$$

donde a, b, c, d, k_1 y k_2 son constantes. Nos preguntamos cuándo podemos resolver el sistema para u y v en términos de x y y. Si escribimos el sistema como

$$\begin{aligned} au + bv &= k_1x \\ cu + dv &= k_2y \end{aligned}$$

podemos ver más claro que la pregunta es: ¿cuándo tiene este sistema solución para las incógnitas u y v? La respuesta a esta pregunta simple la sabemos desde hace mucho: cuando

$$\Delta = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \neq 0$$

En tal caso podemos escribir

$$u = \frac{1}{\Delta}(k_1dx - k_2by), \qquad v = \frac{1}{\Delta}(k_2ay - k_1cx)$$

Esta respuesta no cambiaría si consideráramos el sistema

$$\begin{aligned} au + bv &= f_1(x, y) \\ cu + dv &= f_2(x, y) \end{aligned}$$

donde f_1 y f_2 son funciones dadas de las variables x, y. La posibilidad de despejar las variables u y v en términos de x y y recae sobre los coeficientes de estas variables en las ecuaciones dadas.

Por supuesto, este problema pasa de su gran simplicidad a una posible gran complejidad si ahora las ecuaciones *ya no son lineales en u y v*. Es decir, si ahora escribimos el sistema como

$$\begin{aligned} g_1(u, v) &= f_1(x, y) \\ g_2(u, v) &= f_2(x, y) \end{aligned}$$

en que g_1 y g_2 son ciertas funciones de u y v, nos preguntamos cuándo de él podemos despejar a u y v en términos de x y y. Más generalmente, consideremos el problema siguiente: dadas las funciones F y G de las variables u, v, x, y, nos preguntamos cuándo de las expresiones

$$\left.\begin{aligned} F(x, y, u, v) &= 0 \\ G(x, y, u, v) &= 0 \end{aligned}\right\} \qquad (*)$$

podemos despejar a u y v en términos de x y y. En el caso de que esto sea posible diremos que las funciones $u = \varphi_1(x, y)$ y $v = \varphi_2(x, y)$ son *funciones implícitas* dadas en el par de ecuaciones (*).

Por la experiencia de la sección anterior, no debe resultarnos extraño el hecho de que la respuesta a la pregunta planteada tiene sólo respuesta *local*. Es decir, si $\mathbf{p} = (\bar{x}, \bar{y}, \bar{u}, \bar{v})$ es un punto tal que $F(\mathbf{p}) = G(\mathbf{p}) = 0$ y F y G tienen ciertas propiedades en sus derivadas parciales en los alrededores de $\mathbf{p}$, se podrá esperar la existencia de funciones $u = \varphi_1(x, y)$, $v = \varphi_2(x, y)$ tales que $\bar{u} = \varphi_1(\bar{x}, \bar{y})$, $\bar{v} = \varphi_2(\bar{u}, \bar{v})$ y

$$F(x, y, \varphi_1(x, y), \varphi_2(x, y)) \equiv 0$$
$$G(x, y, \varphi_1(x, y), \varphi_2(x, y)) \equiv 0$$

con (x, y) en alguna vecindad V de $(\bar{x}, \bar{y})$, y con las funciones φ_1 y φ_2 con ciertas propiedades de diferenciabilidad en V.

Asumamos por el momento la existencia de las funciones φ_1 y φ_2 y veamos cuáles tendrían que ser sus derivadas parciales. Derivando las expresiones (*) respecto de x, recordando que $u = \varphi_1(x, y)$ y $v = \varphi_2(x, y)$, se tiene, aplicando la regla de la cadena

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} = 0$$
$$\frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} = 0$$

que se pueden reescribir como

$$\frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} = -\frac{\partial F}{\partial x}$$
$$\frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} = -\frac{\partial G}{\partial x}$$

y verse como un sistema de dos ecuaciones con dos incógnitas $\frac{\partial u}{\partial x}$ y $\frac{\partial v}{\partial x}$. Se ve entonces que una condición que debe cumplir F y G es, para que este sistema tenga solución, que

$$\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix} \neq 0$$

(en $\mathbf{p}$). En tal caso, según la regla de la Cramer, podemos despejar las derivadas que nos interesan quedándonos

$$\frac{\partial u}{\partial x} = \frac{\det \begin{bmatrix} -\frac{\partial F}{\partial x} & \frac{\partial F}{\partial v} \\ -\frac{\partial G}{\partial x} & \frac{\partial G}{\partial v} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix}}, \qquad \frac{\partial v}{\partial x} = \frac{\det \begin{bmatrix} \frac{\partial F}{\partial u} & -\frac{\partial F}{\partial x} \\ \frac{\partial G}{\partial u} & -\frac{\partial G}{\partial x} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix}}$$

En forma análoga, si derivamos la expresión (*) respecto de y obtenemos

$$\frac{\partial F}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial y} = -\frac{\partial F}{\partial y}$$
$$\frac{\partial G}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial y} = -\frac{\partial G}{\partial y}$$

de donde

$$\frac{\partial u}{\partial y} = \frac{\det\begin{bmatrix} -\dfrac{\partial F}{\partial y} & \dfrac{\partial F}{\partial v} \\ -\dfrac{\partial G}{\partial y} & \dfrac{\partial G}{\partial v} \end{bmatrix}}{\det\begin{bmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial G}{\partial u} \\ \dfrac{\partial F}{\partial v} & \dfrac{\partial G}{\partial v} \end{bmatrix}}, \qquad \frac{\partial v}{\partial y} = \frac{\det\begin{bmatrix} \dfrac{\partial F}{\partial u} & -\dfrac{\partial F}{\partial y} \\ \dfrac{\partial G}{\partial u} & -\dfrac{\partial G}{\partial y} \end{bmatrix}}{\det\begin{bmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{bmatrix}}$$

Introduzcamos la siguiente notación: Si X, Y son funciones de las variables x, y se llama *jacobiano* de X y Y respecto de x y y, denotado por

$$J\left(\frac{X, Y}{x, y}\right) \quad \text{o} \quad \frac{\partial(X, Y)}{\partial(x, y)}$$

al determinante

$$\frac{\partial(X, Y)}{\partial(x, y)} = \det\begin{bmatrix} \dfrac{\partial X}{\partial x} & \dfrac{\partial X}{\partial y} \\ \dfrac{\partial Y}{\partial x} & \dfrac{\partial Y}{\partial y} \end{bmatrix}$$

Con esta notación las derivadas parciales de $u = \varphi_1(x, y)$, $v = \varphi_2(x, y)$ determinadas anteriormente se ven como

$$(**) \quad \begin{cases} \dfrac{\partial u}{\partial x} = -\dfrac{\dfrac{\partial(F, G)}{\partial(x, v)}}{\dfrac{\partial(F, G)}{\partial(u, v)}} & \dfrac{\partial v}{\partial x} = -\dfrac{\dfrac{\partial(F, G)}{\partial(u, x)}}{\dfrac{\partial(F, G)}{\partial(u, v)}} \\[2ex] \dfrac{\partial u}{\partial y} = -\dfrac{\dfrac{\partial(F, G)}{\partial(y, v)}}{\dfrac{\partial(F, G)}{\partial(u, v)}} & \dfrac{\partial v}{\partial y} = -\dfrac{\dfrac{\partial(F, G)}{\partial(u, y)}}{\dfrac{\partial(F, G)}{\partial(u, v)}} \end{cases}$$

Obsérvese que la condición aquí impuesta de que el jacobiano $\frac{\partial(F,G)}{\partial(u,v)}$ sea no nulo (en $\mathbf{p}$), para nuestro sistema lineal que consideramos al comienzo de la sección

$$F(x, y, u, v) = au + bv - f_1(x, y) = 0$$
$$G(x, y, u, v) = cu + dv - f_2(x, y) = 0$$

se ve como

$$\frac{\partial(F, G)}{\partial(u, v)} = \det\begin{bmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{bmatrix} = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} \neq 0$$

como ya lo habíamos dicho antes.

Enunciamos ahora el TFIm en este caso de dos expresiones en cuatro variables que definen implícitamente a dos de ellas como funciones de las dos restantes.

Teorema 3.5.1 (De la función implícita tercera versión). Considere las funciones $z_1 = F(x, y, u, v)$, $z_2 = G(x, y, u, v)$. Sea $\mathbf{p} = (\bar{x}, \bar{y}, \bar{u}, \bar{v}) \in \mathbb{R}^4$ un punto tal que $F(\mathbf{p}) = G(\mathbf{p}) = 0$. Suponga que en una bola B (en $\mathbb{R}^4$) de centro en $\mathbf{p}$, las funciones F y G tienen (sus cuatro) derivadas parciales continuas. Si el jacobiano $\frac{\partial(F,G)}{\partial(u,v)}(\mathbf{p}) \neq 0$, entonces las expresiones $F(x, y, u, v) = 0$ y $G(x, y, u, v) = 0$ definen funciones (implícitas) $u = \varphi_1(x, y)$, $v = \varphi_2(x, y)$ definidas en una vecindad V de $(\bar{x}, \bar{y})$, las cuales tienen derivadas parciales continuas en V que se pueden calcular con las fórmulas (**). ■

La demostración de este teorema se basa por completo en el TFIm segunda versión, establecido en la sección anterior (Teorema 3.4.2) y resulta un ejercicio ilustrativo cómo se puede usar tal teorema para probar otros resultados. De cualquier modo, presentamos como opcional su demostración.

Demostración. (Opcional). Como, por hipótesis,

$$\frac{\partial(F,G)}{\partial(u,v)} = \det \begin{bmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{bmatrix}_{\mathbf{p}} \neq 0$$

los cuatro elementos $\frac{\partial F}{\partial u}(\mathbf{p})$, $\frac{\partial F}{\partial v}(\mathbf{p})$, $\frac{\partial G}{\partial u}(\mathbf{p})$, $\frac{\partial G}{\partial v}(\mathbf{p})$ no pueden ser simultáneamente cero. Sin pérdida de generalidad podemos suponer que $\frac{\partial G}{\partial v}(\mathbf{p}) \neq 0$. Entonces, la función $z_1 = G(x, y, u, v)$ satisface las hipótesis del Teorema 3.4.2, del cual podemos concluir que en una bola B de $\mathbb{R}^4$ con centro en $\mathbf{p}$, se puede escribir v como función de x, y, u. Sea $v = \psi(x, y, u)$ tal función y consideremos la función

$$H(x, y, u) = F(x, y, u, \psi(x, y, u))$$

Tenemos que, derivando esta expresión respecto de u

$$\frac{\partial H}{\partial x} = \frac{\partial F}{\partial u} + \frac{\partial F}{\partial v}\frac{\partial \psi}{\partial u}$$

El mismo teorema 3.4.2 nos dice que

$$\frac{\partial \psi}{\partial u} = -\frac{\dfrac{\partial G}{\partial u}}{\dfrac{\partial G}{\partial v}}$$

de modo que

$$\frac{\partial H}{\partial u} = \frac{\partial F}{\partial u} + \frac{\partial F}{\partial v}\left(-\frac{\dfrac{\partial G}{\partial u}}{\dfrac{\partial G}{\partial v}}\right) = \frac{\dfrac{\partial F}{\partial u}\dfrac{\partial G}{\partial v} - \dfrac{\partial F}{\partial v}\dfrac{\partial G}{\partial u}}{\dfrac{\partial G}{\partial v}}$$

es decir que la función $H(x, y, u)$ en el punto $\mathbf{p}$ es tal que

$$\frac{\partial H}{\partial u} = \frac{\dfrac{\partial(F,G)}{\partial(u,v)}(\mathbf{p})}{\dfrac{\partial G}{\partial v}(\mathbf{p})} \neq 0$$

Nuevamente el teorema 3.4.2 nos permite concluir que de la expresión $H(x, y, u) = 0$, podemos, localmente, ver a u como función de x, y. Sea $u = \varphi_1(x, y)$ tal función. Entonces $u = \varphi_1(x, y)$ y

$$v = \psi(x, y, u) = \psi(x, y, \varphi_1(x, y)) = \varphi_2(x, y)$$

Estas son las funciones cuya existencia asegura el teorema. Una vez concluida la existencia de estas funciones es posible obtener las fórmulas (**) para sus derivadas parciales procediendo como se hizo en el libro, o bien, haciendo uso repetido (y cuidadoso) de las fórmulas del teorema 3.4.2. Este es un bonito ejercicio de "sustitución" de expresiones con derivadas, para el lector. Q.E.D.

Ejemplo 1. Considere las expresiones

$$F(x, y, u, v) = xe^{u+v} + uv - 1 = 0$$
$$G(x, y, u, v) = ye^{u-v} - 2uv - 1 = 0$$

En el punto $\mathbf{p} = (1, 1, 0, 0)$ se tiene $F(\mathbf{p}) = G(\mathbf{p}) = 0$. Las derivadas parciales de F y G son

$$\frac{\partial F}{\partial x} = e^{u+v}, \quad \frac{\partial F}{\partial y} = 0, \quad \frac{\partial F}{\partial u} = xe^{u+v} + v, \quad \frac{\partial F}{\partial v} = xe^{u+v} + u$$
$$\frac{\partial G}{\partial x} = 0, \quad \frac{\partial G}{\partial y} = e^{u-v}, \quad \frac{\partial G}{\partial u} = ye^{u-v} - 2u, \quad \frac{\partial G}{\partial v} = -ye^{u-v} - 2u$$

las cuales siempre son continuas. El jacobiano $\frac{\partial(F,G)}{\partial(u,v)}$ es

$$\frac{\partial(F, G)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix} = \det \begin{bmatrix} xe^{u+v} + u & xe^{u+v} + v \\ ye^{u-v} - 2v & -ye^{u-v} - 2u \end{bmatrix}$$

que en el punto $\mathbf{p}$ vale

$$\frac{\partial(F, G)}{\partial(u, v)}(\mathbf{p}) = \det \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = -2 \neq 0$$

Entonces el TFIm nos asegura que en los alrededores de $\mathbf{p}$ podemos (teóricamente) despejar a u y v en términos de x y y y establecer así funciones implícitas $u = u(x, y)$, $v = v(x, y)$, las cuales, aun sin poder hacerlas explícitas, podemos obtener sus derivadas parciales en una vecindad de $(1, 1)$. Por ejemplo

$$\frac{\partial u}{\partial x} = -\frac{\frac{\partial(F, G)}{\partial(x, v)}}{\frac{\partial(F, G)}{\partial(u, v)}} = -\frac{\det \begin{bmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial v} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix}}$$

$$= -\frac{\det\begin{bmatrix} e^{u+v} & xe^{u+v}+v \\ 0 & -ye^{u-v}-2u \end{bmatrix}}{\det\begin{bmatrix} xe^{u+v}+u & xe^{u+v}+v \\ ye^{u-v}-2v & -ye^{u-v}-2u \end{bmatrix}}$$

$$= -\frac{ye^{2u}+2ue^{u+v}}{2xye^{2u}+2(u-v)xe^{u+v}+(u+v)ye^{u-v}}$$ ■

Ejemplo 2. La función $z = f(x, y)$ está dada como $z = u+v$, donde u y v son funciones implícitas determinadas por el par de expresiones

$$F(x, y, u, v) = u + e^{u+v} - x = 0$$
$$G(x, y, u, v) = v + e^{u-v} - y = 0$$

Se quiere hallar la ecuación del plano tangente a $z = f(x, y)$ en el punto correspondiente a $x = y = 1$, $u = v = 0$ (en el que $\frac{\partial(F,G)}{\partial(u,v)} = -1 \neq 0$ y por lo tanto, existen tales funciones u y v).

Necesitamos entonces calcular

$$\frac{\partial z}{\partial x}(1, 1) \qquad \text{y} \qquad \frac{\partial z}{\partial y}(1, 1)$$

Se tiene

$$\frac{\partial z}{\partial x} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x}, \qquad \frac{\partial z}{\partial y} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}$$

Calculemos entonces $\frac{\partial u}{\partial x}(1, 1)$, $\frac{\partial u}{\partial y}(1, 1)$, $\frac{\partial v}{\partial x}(1, 1)$, $\frac{\partial v}{\partial y}(1, 1)$. Estas derivadas se pueden calcular usando las fórmulas establecidas con los jacobianos, o bien, derivando directamente las expresiones $F = 0$ y $G = 0$. Procediendo de esta segunda manera (que en general resulta ser más práctica cuando se procuran resultados numéricos), derivamos respecto de x las expresiones $F = 0$ y $G = 0$ (recordando que u y v son funciones de x y y)

$$\frac{\partial u}{\partial x} + e^{u+v}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial x}\right) - 1 = 0 \qquad \frac{\partial v}{\partial x} + e^{u-v}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial x}\right) = 0$$

de modo que evaluando estas expresiones en los valores $x = y = 1$, $u = v = 0$, obtenemos

$$2\frac{\partial u}{\partial x} + \frac{\partial v}{\partial x} = 1 \qquad \frac{\partial u}{\partial x} = 0$$

de donde $\frac{\partial u}{\partial x} = 0$, $\frac{\partial v}{\partial x} = 1$. Igualmente, al derivar respecto de y las expresiones $F = 0$, $G = 0$ obtenemos

$$\frac{\partial u}{\partial y} + e^{u+v}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}\right) = 0$$
$$\frac{\partial v}{\partial y} + e^{u-v}\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial y}\right) - 1 = 0$$

de donde, al evaluar con $x = y = 1$, $u = v = 0$ obtenemos $\frac{\partial u}{\partial y} = 1$, $\frac{\partial v}{\partial y} = -2$ Así pues

$$\frac{\partial z}{\partial x} = 0 + 1 = 1, \qquad \frac{\partial z}{\partial y} = 1 + (-2) = -1$$

de modo que la ecuación del plano tangente procurado es

$$z = x - 1 - (y - 1)$$

o sea

$$z = x - y$$

■

Ejemplo 3. Dado un par de expresiones $F(x, y, u, v) = 0$, $G(x, y, u, v) = 0$ que determinan (localmente, alrededor de algún punto $\mathbf{p}$) funciones $u = \varphi_1(x, y)$, $v = \varphi_1(x, y)$, nos podríamos preguntar (con F y G suficientemente bien portadas) por derivadas de segundo orden de estas funciones. La discusión general nos llevará a expresiones bastante complicadas para estas derivadas y no es nuestra intención obtenerlas. Sin embargo, sí puede ser un buen ejercicio de derivación y uso de la regla de la cadena hacer esto en un caso particular. Considere las expresiones

$$F(x, y, u, v) = u - v + 2x - 2y = 0$$
$$G(x, y, u, v) = 3u^3 + v^3 - 5x^2 + y^3 = 0$$

Como

$$\frac{\partial(F, G)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix} = \det \begin{bmatrix} 1 & -1 \\ 9u^2 & 3v^2 \end{bmatrix} = 3v^2 + 9u^2 = 0 \Leftrightarrow u = v = 0$$

consideremos los alrededores de cualquier punto $(x, y, u, v) \in \mathbb{R}^4$ con u y v no nulos, por ejemplo, el punto $(1, 1, 1, 1)$. Se quiere calcular $\frac{\partial^2 u}{\partial y^2}$ en los alrededores de este punto. Tenemos

$$\frac{\partial u}{\partial y} = -\frac{\dfrac{\partial(F, G)}{\partial(y, v)}}{\dfrac{\partial(F, G)}{\partial(u, v)}} = -\frac{1}{3v^2 + 9u^2} \det \begin{bmatrix} \frac{\partial F}{\partial y} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial y} & \frac{\partial G}{\partial v} \end{bmatrix}$$
$$= -\frac{1}{3v^2 + 9u^2} \det \begin{bmatrix} -2 & -1 \\ 3y^2 & 3v^2 \end{bmatrix} = \frac{6v^2 - 3y^2}{3v^2 + 9u^2} = \frac{2v^2 - y^2}{v^2 + 3u^2}$$

Ahora

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{2v^2 - y^2}{v^2 + 3u^2}\right)$$
$$= \frac{(v^2 + 3u^2)\left(4v\dfrac{\partial u}{\partial y} - 2y\right) - (2v^2 - y^2)\left(2v\dfrac{\partial v}{\partial y} + 6u\dfrac{\partial u}{\partial y}\right)}{(v^2 + 3u^2)^2}$$
$$= \frac{2v(y^2 + 6u^2)\dfrac{\partial v}{\partial y} - (2v^2 - y^2)6u\dfrac{\partial u}{\partial y} - (v^2 + 3u^2)2y}{(v^2 + 3u^2)^2}$$

Calculemos $\frac{\partial v}{\partial y}$

$$\frac{\partial v}{\partial y} = -\frac{\frac{\partial(F,G)}{\partial(u,y)}}{\frac{\partial(F,G)}{\partial(u,v)}} = -\frac{1}{3v^2+9u^2}\det\begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial y} \end{bmatrix}$$

$$= -\frac{1}{3v^2+9u^2}\det\begin{bmatrix} 1 & -2 \\ 9u^2 & 3y^2 \end{bmatrix} = -\frac{3y^2+18u^2}{3v^2+9u^2} = -\frac{y^2+6u^2}{v^2+3u^2}$$

sustituyendo esta expresión junto con la de $\frac{\partial u}{\partial y}$ en la expresión para $\frac{\partial^2 u}{\partial y^2}$ nos queda finalmente

$$\frac{\partial^2 u}{\partial y^2} = -\frac{2v(y^2+6u^2)^2 + 6u(2v^2-y^2)^2 - 2y(v^2+3u^2)^2}{(v^2+3u^2)^3}$$ ■

Presentamos por último la versión del TFIm más general que consideramos en este texto: cuando se tienen n expresiones del tipo $F_i(x_1, x_2, \ldots, x_m, y_1, \ldots, y_n) = 0$, $i = 1, 2, \ldots, n$ con $m + n$ variables, las cuales definen a n de ellas, digamos $y_1, \ldots, y_n$ como funciones implícitas de las m restantes, $x_1, x_2, \ldots, x_m$.

Teorema 3.5.2 (De la Función Implícita. Cuarta versión). Considere las n funciones

$$u_i = F_i(x_1, \ldots, x_m, y_1, \ldots, y_n), \qquad i = 1, 2, \ldots, n$$

Sea $\mathbf{p} = (\bar{x}_1, \ldots, \bar{x}_m, \bar{y}_1, \ldots, \bar{y}_n) \in \mathbb{R}^{m+n}$ un punto tal que $F_i(\mathbf{p}) = 0$, $i = 1, 2, \ldots, n$. Suponga que en una bola B de $\mathbb{R}^{m+n}$ con centro en $\mathbf{p}$, las funciones F_i tienen (sus $m + n$) derivadas parciales continuas. Si el jacobiano

$$\frac{\partial(F_1, F_2, \ldots, F_n)}{\partial(y_1, y_2, \ldots, y_n)} = \det\begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} & \cdots & \frac{\partial F_2}{\partial y_n} \\ \frac{\partial F_n}{\partial y_1} & \frac{\partial F_n}{\partial y_2} & \cdots & \frac{\partial F_n}{\partial y_n} \end{bmatrix}$$

es no nulo en $\mathbf{p}$, entonces las expresiones $F_i(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0$ definen funciones (implícitas) $y_i = \varphi_i(x_1, \ldots, x_m)$, $i = 1, 2, \ldots, n$ definidas en una vecindad V de $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m)$, las cuales tiene derivadas parciales continuas en V que se pueden calcular como

$$\frac{\partial y_i}{\partial x_j} = -\frac{\frac{\partial(F_1, F_2, \ldots, F_n)}{\partial(y_1, \ldots, y_{i-1}, x_j, y_{i+1}, \ldots, y_n)}}{\frac{\partial(F_1, F_2, \ldots, F_n)}{\partial(y_1, y_2, \ldots, y_n)}}$$ ■

Ejemplo 4. Considere las expresiones

$$F(x, y, u, v, w) = x + y + u + v + w = 0$$
$$G(x, y, u, v, w) = x^2 - y^2 + u^2 - 2v^2 + w^2 + 1 = 0$$
$$H(x, y, u, v, w) = x^3 + y^3 + u^4 - 3v^4 + 8w^4 + 2 = 0$$

En el punto $\mathbf{p} = (1, -1, 1, -1, 0)$, se tiene $F(\mathbf{p}) = G(\mathbf{p}) = H(\mathbf{p}) = 0$. Todas las derivadas parciales de F, G y H son continuas siempre. Se tiene además

$$\frac{\partial(F, G, H)}{\partial(u, v, w)}(\mathbf{p}) = \det \begin{bmatrix} 1 & 1 & 1 \\ 2u & -4v & 2w \\ 4u^3 & -12v^3 & 32w^2 \end{bmatrix}_{\substack{u=1 \\ v=-1 \\ w=0}} = 8 \neq 0$$

El Teorema anterior nos asegura entonces que en torno a $\mathbf{p}$ podemos despejar u, v, w en términos de x, y y establecer así funciones $u = u(x, y)$, $v = v(x, y)$, $w = w(x, y)$, las cuales, aun cuando no se puedan ver explícitamente, sabemos que tienen derivadas continuas en una vecindad de $(1, -1)$. Calculemos, por ejemplo, $\frac{\partial v}{\partial y}(1, -1)$. Como

$$\frac{\partial v}{\partial y} = -\frac{\dfrac{\partial(F, G, H)}{\partial(u, y, w)}}{\dfrac{\partial(F, G, H)}{\partial(u, v, w)}}$$

se tiene

$$\frac{\partial(F, G, H)}{\partial(u, y, w)} = \det \begin{bmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial y} & \dfrac{\partial F}{\partial w} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial y} & \dfrac{\partial G}{\partial w} \\ \dfrac{\partial H}{\partial u} & \dfrac{\partial H}{\partial y} & \dfrac{\partial H}{\partial w} \end{bmatrix} = \det \begin{bmatrix} 1 & 1 & 1 \\ 2u & -2y & 2w \\ 4u^3 & 3y^2 & 32w^3 \end{bmatrix}$$

Al evaluar en $\mathbf{p}$ obtenemos

$$\frac{\partial(F, G, H)}{\partial(u, v, w)} = \det \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 0 \\ 4 & 3 & 0 \end{bmatrix} = -2$$

Así,

$$\frac{\partial v}{\partial y}(1, -1) = -\frac{-2}{8} = \frac{1}{4}$$

■

Ejercicios (Capítulo 3, Sección 5)

1. El sistema

$$u - v = x + y, \quad u + v = x - y$$

define funciones implícitas $u = u(x, y)$, $v = v(x, y)$, $x = x(u, v)$, $y = y(u, v)$, las cuales se pueden hacer explícitas. Obtenga las derivadas parciales de estas funciones. Compruebe que

$$\frac{\partial(u, v)}{\partial(x, y)} \frac{\partial(x, y)}{\partial(u, v)} = 1$$

2. Considere las funciones $u = u(x, y)$, $v = v(x, y)$ definidas implícitamente por las expresiones

$$e^u + e^v = x + ye, \quad ue^u + ve^v = xye$$

Calcule las derivadas parciales $\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial y}$, para $u = 0$, $v = 1$, $x = 1$, $y = 1$.

3. Considere las expresiones

$$uv - 3x + 2y = 0, \qquad u^4 - v^4 = x^2 - y^2$$

Habiendo verificado que éstas definen funciones $u = u(x, y)$, $v = v(x, y)$ en los alrededores del punto $(u, v, x, y) = (1, 1, 1, 1)$, determine las ecuaciones de los planos tangentes a las superficies $u = u(x, y)$, $v = v(x, y)$ en $\mathbf{p}$.

4. Sean $F, G\colon \mathbb{R}^4 \to \mathbb{R}$ dos funciones de clase $\mathscr{C}^1$ tales que $\operatorname{grad} F(1, 1, 1, 1) = (3, 2, 1, -1)$, $\operatorname{grad} G(1, 1, 1, 1) = (4, -5, 2, 2)$. Suponga que las expresiones $F(x, y, u, v) = 0$, $G(x, y, u, v) = 0$ determinan funciones de clase $\mathscr{C}^1$, $u = u(x, y)$, $v = v(x, y)$ alrededor del punto $\mathbf{p} = (1, 1, 1, 1)$. Demuestre que las expresiones $F(x^2, y^2, u, v) = 0$, $G(x^3, y^4, u, v) = 0$ determinan también funciones de clase $\mathscr{C}^1$, $\tilde{u} = \tilde{u}(x, y)$, $\tilde{v} = \tilde{v}(x, y)$ en torno a $\mathbf{p}$. Calcule las derivadas parciales de estas funciones para $x = y = 1$.

5. Sean $F, G\colon \mathbb{R}^3 \to \mathbb{R}$ dos funciones de clase $\mathscr{C}^1$. Considere el sistema

$$F(x, y, z) = 0, \quad G(x, y, z) = 0$$

Sea $\mathbf{p} = (x_0, y_0, z_0)$ un punto en el que $F(\mathbf{p}) = G(\mathbf{p}) = 0$. Establezca condiciones bajo las cuales estas expresiones determinen funciones implícitas $x = x(z)$, $y = y(z)$. En tal caso, halle $x'(z_0)$, $y'(z_0)$. Por lo general, establezca condiciones bajo las cuales estas expresiones determinen *algunas* de las funciones implícitas $x = x(z)$, $y = y(z)$, o, $x = x(y)$, $z = z(y)$, o, $y = y(x)$, $z = z(x)$.

6. Como caso particular del ejercicio anterior, considere las funciones $F(x, y, z) = a_1x + b_1y + c_1z + d_1$, $G(x, y, z) = a_2x + b_2y + c_2z + d_2$. Verifique que las condiciones establecidas en el ejercicio anterior para que de las expresiones $F(x, y, z) = 0$, $G(x, y, z) = 0$ se puedan despejar *dos de las variables x, y, z en términos de la restante*, se ven como

1. $a_1b_2 - a_2b_1 \neq 0$, o, **2.** $a_1c_2 - a_2c_1 \neq 0$, o, **3.** $b_1c_2 - b_2c_1 \neq 0$

(correspondiendo a los casos de existencia de funciones: **1.** $x = x(z)$, $y = y(z)$; **2.** $x = x(y)$, $z = z(y)$; **3.** $y = y(x)$, $z = z(x)$, respectivamente). Constate que estas condiciones son equivalentes a la independencia lineal de los vectores $v_1 = (a_1, b_1, c_1)$, $v_2 = (a_2, b_2, c_2)$, que son los vectores normales a los planos que representan $F(x, y, z) = 0$, $G(x, y, z) = 0$. Relacione el resultado aquí establecido con las condiciones para que dos planos en $\mathbb{R}^3$ se corten entre sí (en una línea recta).

7. Considere las expresiones $F(x, y, z) = 0$, $G(x, y, z) = 0$. Sea $\mathbf{p} \in \mathbb{R}^3$ un punto para el que $F(\mathbf{p}) = G(\mathbf{p}) = 0$. Suponga que estas expresiones determinan, localmente, en una bola $B(\mathbf{p})$ del punto $\mathbf{p}$, funciones $x = x(z)$, $y = y(z)$. En tal caso, el conjunto de puntos $(x, y, z) \in \mathbb{R}^3$ tales que $x = x(t)$, $y = y(t)$, $z = t$, con $t \in \mathbb{R}$ tal que $(x(t), y(t), t) \in B(\mathbf{p})$, puede verse como

el conjunto de puntos (x, y, z) *de la curva de intersección de las dos superficies* $F(x, y, z) = 0$, $G(x, y, z) = 0$. Observe que esta curva podría ser descrita (en su caso) como

$$C = \{(x, y, z)|x = x(t), y = t, z = z(t)\}$$

o

$$C = \{(x, y, z)|x = t, y = y(t), z = z(t)\}$$

con $t \in \mathbb{R}$ tal que $(x(t), t, z(t)) \in B(\mathbf{p})$ o $(t, x(t), y(t)) \in B(\mathbf{p})$, respectivamente. ¿A qué casos corresponden estas descripciones de la curva C? Considere, por ejemplo, las superficies

$$F(x, y, z) = x^2 + y^2 + z^2 - 2x - 1 = 0$$
$$G(x, y, z) = x^2 + y^2 + z^2 - 2y - 1 = 0$$

y el punto $\mathbf{p} = (1, 1, 1)$. Compruebe que $\frac{\partial(F,G)}{\partial(x,y)}(\mathbf{p}) = -4$. Concluya entonces que en los alrededores de $\mathbf{p}$ podemos ver la curva de intersección de estas dos superficies como $x = x(t)$, $y = y(t)$, $z = t$. Halle las funciones $x(t)$, $y(t)$. Describa geométricamente esta situación.

8. Considere las expresiones

$$x \operatorname{sen} y + y \operatorname{sen} x = u \operatorname{sen} v + v \operatorname{sen} u$$
$$x \cos y + y \cos x = u \cos v + v \cos u$$

Compruebe que en los alrededores del punto $\mathbf{p} = (x, y, u, v) = (0, 1, 0, 1)$ estas expresiones determinan funciones de clase $\mathscr{C}^1$, $u = u(x, y)$, $v = v(x, y)$, $x = x(u, v)$, $y = y(u, v)$. Halle las derivadas parciales de estas funciones en el punto $\mathbf{p}$. Considere luego las funciones Φ, $\Psi: \mathbb{R}^2 \to \mathbb{R}^2$, $\Phi(x, y) = (u(x, y), v(x, y))$, $\Psi(u, v) = (x(u, v), y(u, v))$. Escriba las matrices jacobianas $J\Phi(0, 1)$, $J\Psi(0, 1)$. Determine los productos $J\Phi(0, 1)J\Psi(0, 1)$ y $J\Psi(0, 1)J\Phi(0, 1)$.

En los ejercicios 9–15 establezca condiciones bajo las cuales las expresiones dadas determinan funciones de clase $\mathscr{C}^1$ $u = u(x, y)$, $v = v(x, y)$. En cada caso, halle la expresión de la derivada parcial indicada, en términos de las (derivadas parciales de las) funciones dadas f, $g: \mathbb{R}^2 \to \mathbb{R}$, las cuales se suponen de clase $\mathscr{C}^1$.

9. $f(x, y) = g(u, v)$, $\quad g(x, y) = f(u, v)$, $\quad \dfrac{\partial u}{\partial x} = ?$

10. $f(x, y) = g(u, v)$, $\quad f(y, x) = g(v, u)$, $\quad \dfrac{\partial v}{\partial x} = ?$

11. $f(x, u) = g(y, v)$, $\quad f(u, x) = g(v, y)$, $\quad \dfrac{\partial u}{\partial y} = ?$

12. $f(x, y) = f(u, v)$, $\quad g(x, y) = g(u, v)$, $\quad \dfrac{\partial u}{\partial y} = ?$

13. $yf(x, x) = ug(v, v)$, $xf(y, y) = vg(u, u)$, $\quad \dfrac{\partial u}{\partial x} = ?$

14. $xuf(y, v) = g(u, v)$, $yvg(x, u) = f(u, v)$, $\quad \dfrac{\partial v}{\partial x} = ?$

15. $f(xu, yv) = g(u, v)$, $f(yu, xv) = g(u, v)$, $\dfrac{\partial u}{\partial x} = ?$

16. Considere las expresiones

$$x + y = e^u - e^v, \quad x^2 + y^2 = u + v$$

Demuestre que estas expresiones determinan funciones implícitas $u = u(x, y)$, $v = v(x, y)$ en los alrededores del punto $\mathbf{p} = (0, 0, 0, 0)$. Halle las derivadas parciales de segundo orden cruzadas de estas funciones en el origen.

17. Sean $f_1, f_2, g_1, g_2: \mathbb{R} \to \mathbb{R}$ funciones continuas tales que $f_1(1) = f_2(1) = g_1(1) = g_2(1) = 1$. Considere las expresiones

$$\int_u^{v^2} f_1(t)dt = \int_x^y g_1(t)dt, \qquad \int_{u^3}^{v^4} f_2(t)dt = \int_{x^2}^{y^2} g_2(t)dt$$

Demuestre que éstas determinan funciones implícitas $u = u(x, y)$, $v = v(x, y)$ en los alrededores del punto $\mathbf{p} = (1, 1, 1, 1)$. Determine las derivadas parciales $\frac{\partial u}{\partial x}(1, 1)$, $\frac{\partial u}{\partial y}(1, 1)$, $\frac{\partial v}{\partial x}(1, 1)$, $\frac{\partial v}{\partial y}(1, 1)$

18. Demuestre que las expresiones

$$3xy + x^2z^2 + z^3 + 2uv - 4v - 3u = 0$$
$$x^3 + y^3 + z^3 - 2v^2 - u = 0$$

determinan funciones implícitas $u = u(x, y, z)$, $v = v(x, y, z)$ en los alrededores del punto $\mathbf{p} = (1, 1, 1, 1, 1)$. Halle las derivadas parciales de estas funciones en el punto $(1, 1, 1)$.

19. Sean $F, G, H: \mathbb{R}^6 \to \mathbb{R}$ funciones de clase $\mathscr{C}^1$ tales que en el punto $\mathbf{p} \in \mathbb{R}^6$ tienen por vectores gradientes a $\operatorname{grad} F(\mathbf{p}) = (1, 0, -1, 1, 1, 2)$, $\operatorname{grad} G(\mathbf{p}) = (0, 0, -2, 0, 1, 1)$, $\operatorname{grad} H(\mathbf{p}) = (2, -2, 0, 0, 0, 1)$. Demuestre que las expresiones

$$F(x, y, z, u, v, w) = 0, \quad G(x, y, z, u, v, w) = 0, \quad H(x, y, z, u, v, w) = 0$$

definen funciones implícitas $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$ en los alrededores del punto $\mathbf{p}$. Calcule las derivadas parciales $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial y}$, $\frac{\partial w}{\partial z}$.

20. Demuestre que las expresiones

$$3x = u + v + w, \quad x^2 + y^2 = u^2 + v^2, \quad x^3 + y^3 + z^3 = 3u^3$$

definen funciones implícitas $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$ alrededor del punto $\mathbf{p} = (1, 1, 1, 1, 1, 1)$. Determine las derivadas parciales $\frac{\partial u}{\partial z}$, $\frac{\partial v}{\partial y}$, $\frac{\partial w}{\partial x}$. Demuestre también que tales expresiones definen funciones implícitas $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$ alrededor del punto $\mathbf{p}$. Calcule las derivadas parciales $\frac{\partial x}{\partial u}$, $\frac{\partial y}{\partial v}$, $\frac{\partial z}{\partial w}$.

En los ejercicios 21–23, sean $f, g, h: \mathbb{R}^3 \to \mathbb{R}$ funciones de clase $\mathscr{C}^1$, cuyos gradientes en el origen de coordenadas son $\operatorname{grad} f(0, 0, 0) = (1, 0, 0)$, $\operatorname{grad} g(0, 0, 0) = (0, 1, 0)$, $\operatorname{grad} h(0, 0, 0) = (0, 0, 1)$. Diga en cada caso si las expresiones dadas determinan funciones implícitas: **a.** $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$; **b.** $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$, en los alrededores del punto $\mathbf{p} = (0, 0, 0, 0, 0, 0)$.

21. $f(x, y, z) = f(u, v, w)$, $g(x, y, z) = g(u, v, w)$, $h(x, y, z) = h(u, v, w)$

22. $f(x, y, z) = f(v, u, w)$, $g(x, y, z) = g(u, v, w)$, $h(x, y, z) = h(u, v, w)$

23. $f(x, x, x) = g(u, u, w)$, $g(y, y, y) = h(u, v, v)$, $h(z, z, z) = f(v, w, w)$

3.6 Funciones inversas

En nuestro primer curso de cálculo se estudió el problema de "invertir" funciones $f: I \subseteq \mathbb{R} \to \mathbb{R}$, en el sentido de que siendo f una función del tipo $y = f(x)$, se quería, a partir de ella obtener la "regla de asociación inversa", $x = f^{-1}(y)$. Esquematicamente

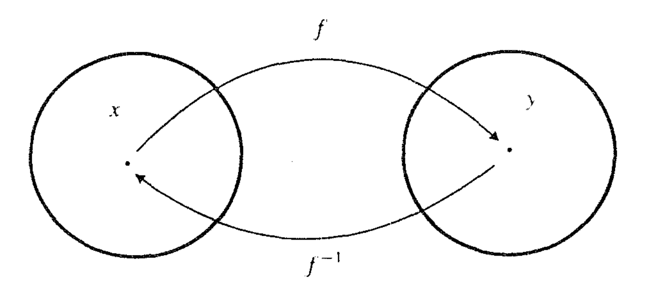


Figura 1. La función inversa.

Para que esta función f^{-1} exista, es claro que f tiene que ser inyectiva, pues en tal caso cada $y_0 = f(x_0)$ está determinada por un único x_0 y entonces se puede definir $x_0 = f^{-1}(y_0)$.

Además se ve también que f^{-1} debe cumplir con $(f^{-1} \circ f)(x) = x$, para toda x en el dominio de f, y, $(f \circ f^{-1})(y) = y$, para toda y en el rango de f.

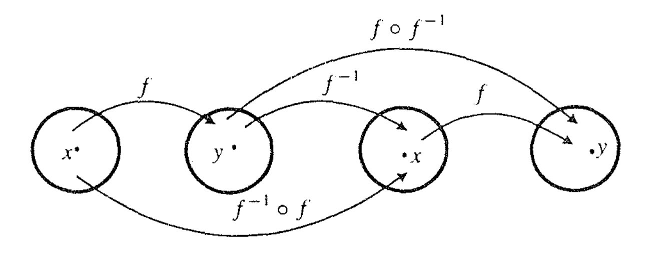


Figura 2. La composición de una función con su inversa es la función identidad.

El resultado más importante relacionado con este tema nos decía que, si para algún punto $x_0 \in$ dominio de f, se tenía $f'(x_0) \neq 0$, entonces, localmente (alrededor de x_0) se podía esperar la existencia de la función f^{-1} la cual era diferenciable y su derivada era "la inversa" (en sentido algebraico) de la derivada de la función f. Con más precisión

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}$$

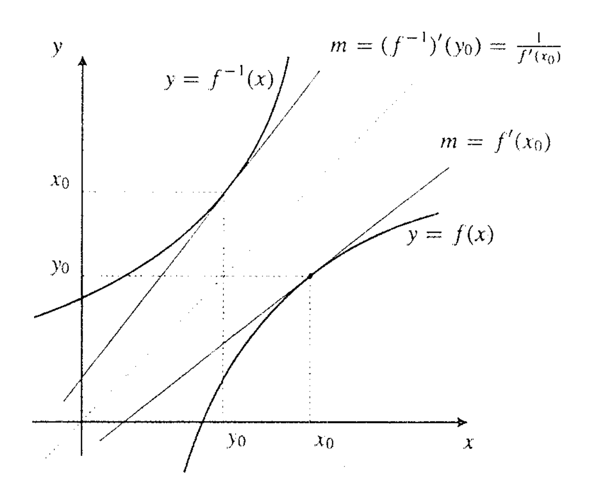


Figura 3. La derivada de la función inversa.

El problema que vamos a considerar ahora será el de estudiar las posibilidades de inversión de una función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, definida en algún conjunto abierto U de $\mathbb{R}^2$. Por supuesto que este estudio será en términos de las propiedades diferenciables de la función F, e interesará también relacionar tales propiedades con las correspondientes a la inversa.

Entendemos, como siempre, por *inversa* de la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ la función $F^{-1}: V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, tal que

$$(F^{-1} \circ F)(u, v) = (u, v), \quad (u, v) \in U$$
$$(F \circ F^{-1})(x, y) = (x, y), \quad (x, y) \in \text{ rango de } F = V$$

escribiremos (u, v) para referirnos a los puntos de $\mathbb{R}^2$ en el dominio de F y (x, y) para los del codominio de F.

Recordando que la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ tiene dos funciones coordenadas $f, g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, digamos $x = f(u, v)$, $y = g(u, v)$, nuestro problema lo podemos ver desde otro punto de vista como: dadas las funciones

$$x = f(u, v)$$
$$y = g(u, v)$$

que describen a x y y como funciones de u y v, se quiere estudiar en qué condiciones es posible establecer "funciones inversas" que describan a u y v como funciones de x y y, digamos

$$u = \varphi(x, y)$$
$$v = \psi(x, y)$$

(estas son las funciones coordenadas de F^{-1}). (figura 4).

Se debe tener entonces que

$$(F^{-1} \circ F)(u, v) = F^{-1}(F(u, v)) = F^{-1}(x, y) = F^{-1}(f(u, v), g(u, v)) = (u, v)$$

y

$$(F \circ F^{-1})(x, y) = F(F^{-1}(x, y)) = F(u, v) = F(\varphi(x, y), \psi(x, y)) = (x, y)$$

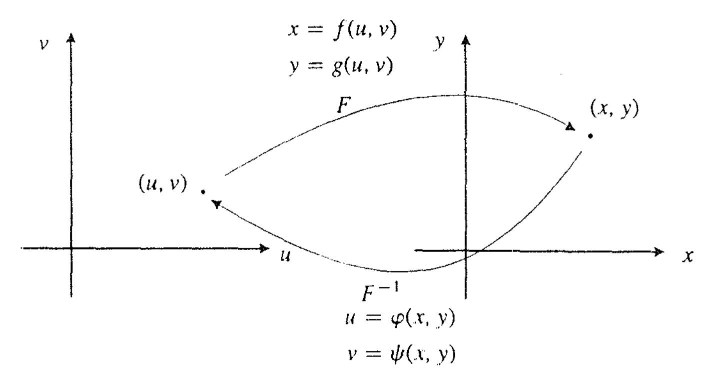


Figura 4. La inversa de F en términos de sus funciones coordenadas.

La ventaja de esta manera de ver el problema es que lo podemos colocar bajo las perspectivas de las funciones implícitas estudiadas en la sección anterior. Más aún, si consideramos las expresiones

$$G(x, y, u, v) = x - f(u, v) = 0$$
$$H(x, y, u, v) = y - g(u, v) = 0$$

lo que pretendemos es "despejar" de ellas a u y v en términos de x y y y establecer así las funciones $u = \varphi(x, y)$, $v = \psi(x, y)$. Entonces, el TFIm (Tercera versión), (Teorema 3.5.1) nos da las condiciones para que podamos hacer esto.

En efecto, aplicando a esta situación el Teorema 3.5.1 tenemos: sea $\mathbf{p} = (\bar{x}, \bar{y}, \bar{u}, \bar{v}) \in \mathbb{R}^4$ un punto tal que $G(\mathbf{p}) = H(\mathbf{p}) = 0$ [lo cual se traduce en nuestro problema como: sea $(\bar{x}, \bar{y}) = F(\bar{u}, \bar{v})$, o bien, sea $\bar{x} = f(\bar{u}, \bar{v})$, $\bar{y} = g(\bar{u}, \bar{v})$]. Supongamos que en una bola B de centro en $\mathbf{p}$ las derivadas parciales de G y H son continuas [lo cual se traduce en nuestro problema como: ya que $\frac{\partial G}{\partial x} = 1$, $\frac{\partial G}{\partial v} = 0$, $\frac{\partial H}{\partial x} = 0$, $\frac{\partial H}{\partial y} = 1$ —que siempre son funciones continuas—, suponiendo que en una bola B —en $\mathbb{R}^3$— de centro en $(\bar{u}, \bar{v})$ las derivadas parciales $\frac{\partial G}{\partial u} = -\frac{\partial f}{\partial u}$, $\frac{\partial G}{\partial v} = -\frac{\partial f}{\partial v}$, $\frac{\partial H}{\partial u} = -\frac{\partial g}{\partial u}$, $\frac{\partial H}{\partial v} = -\frac{\partial g}{\partial v}$ son funciones continuas]. Si el jacobiano $\frac{\partial(G,H)}{\partial(u,v)}$ es no nulo en $\mathbf{p}$ [lo cual se traduce como: si el jacobiano

$$\frac{\partial(G, H)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \\ \frac{\partial H}{\partial u} & \frac{\partial H}{\partial v} \end{bmatrix} = \det \begin{bmatrix} -\frac{\partial f}{\partial u} & -\frac{\partial f}{\partial v} \\ -\frac{\partial g}{\partial u} & -\frac{\partial g}{\partial v} \end{bmatrix}$$
$$= \det \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix} = \frac{\partial(f, g)}{\partial(u, v)}$$

es no nulo en $(\bar{u}, \bar{v})$], entonces es posible despejar u y v en términos de x y y, y establecer así funciones $u = \varphi(x, y)$, $v = \psi(x, y)$ definidas en una vecindad V de $(\bar{x}, \bar{y}) = F(\bar{u}, \bar{v})$, las cuales tienen derivadas parciales continuas en V que se pueden calcular como

$$\frac{\partial u}{\partial x} = -\frac{\frac{\partial(G,H)}{\partial(x,v)}}{\frac{\partial(G,H)}{\partial(u,v)}} = -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} \frac{\partial G}{\partial x} & \frac{\partial G}{\partial v} \\ \frac{\partial H}{\partial x} & \frac{\partial H}{\partial v} \end{bmatrix}$$

$$= -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} 1 & -\frac{\partial f}{\partial v} \\ 0 & -\frac{\partial g}{\partial v} \end{bmatrix} = \frac{\frac{\partial g}{\partial v}}{\frac{\partial(f,g)}{\partial(u,v)}}$$

$$\frac{\partial u}{\partial y} = -\frac{\frac{\partial(G,H)}{\partial(y,v)}}{\frac{\partial(G,H)}{\partial(u,v)}} = -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} \frac{\partial G}{\partial y} & \frac{\partial G}{\partial v} \\ \frac{\partial H}{\partial y} & \frac{\partial H}{\partial v} \end{bmatrix}$$

$$= -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} 0 & -\frac{\partial f}{\partial v} \\ 1 & -\frac{\partial g}{\partial v} \end{bmatrix} = \frac{-\frac{\partial f}{\partial v}}{\frac{\partial(f,g)}{\partial(u,v)}}$$

$$\frac{\partial v}{\partial x} = -\frac{\frac{\partial(G,H)}{\partial(u,x)}}{\frac{\partial(G,H)}{\partial(u,v)}} = -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} \frac{\partial G}{\partial u} & \frac{\partial G}{\partial x} \\ \frac{\partial H}{\partial u} & \frac{\partial H}{\partial x} \end{bmatrix}$$

$$= -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} -\frac{\partial f}{\partial u} & 1 \\ -\frac{\partial g}{\partial u} & 0 \end{bmatrix} = \frac{-\frac{\partial g}{\partial u}}{\frac{\partial(f,g)}{\partial(u,v)}}$$

$$\frac{\partial v}{\partial y} = -\frac{\frac{\partial(G,H)}{\partial(u,y)}}{\frac{\partial(G,H)}{\partial(u,v)}} = -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} \frac{\partial G}{\partial u} & \frac{\partial G}{\partial y} \\ \frac{\partial H}{\partial u} & \frac{\partial H}{\partial y} \end{bmatrix}$$

$$= -\frac{1}{\frac{\partial(f,g)}{\partial(u,v)}} \det \begin{bmatrix} -\frac{\partial f}{\partial u} & 0 \\ -\frac{\partial g}{\partial u} & 1 \end{bmatrix} = \frac{\frac{\partial f}{\partial u}}{\frac{\partial(f,g)}{\partial(u,v)}}$$

En resumen tenemos: sean $f, g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ funciones definidas en el conjunto abierto U de $\mathbb{R}^2$. Sea $\bar{x} = f(\bar{u}, \bar{v})$, $\bar{y} = g(\bar{u}, \bar{v})$. Suponga que en alguna bola B en $\mathbb{R}^2$ de centro $(\bar{u}, \bar{v})$, las derivadas parciales $\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}$ son continuas. Si el jacobiano $\frac{\partial(f,g)}{\partial(u,v)}$ es no nulo en $(\bar{u}, \bar{v})$, entonces existe una vecindad V de $\bar{x}$, $\bar{y}$ donde podemos definir "funciones inversas" $u = \varphi(x, y)$, $v = \psi(x, y)$ (es decir, tales que $\bar{u} = \varphi(\bar{x}, \bar{y})$, $\bar{v} = \psi(\bar{x}, \bar{y})$, y $f(\varphi(x, y), \psi(x, y)) = x$, $g(\varphi(x, y), \psi(x, y)) = y$, para $(x, y) \in V$) las cuales tienen derivadas parciales continuas en V que se calculan como

$$\frac{\partial u}{\partial x} = \frac{\frac{\partial g}{\partial v}}{\frac{\partial(f,g)}{\partial(u,v)}}, \quad \frac{\partial u}{\partial y} = \frac{-\frac{\partial f}{\partial v}}{\frac{\partial(f,g)}{\partial(u,v)}}, \quad \frac{\partial v}{\partial x} = \frac{-\frac{\partial g}{\partial u}}{\frac{\partial(f,g)}{\partial(u,v)}}, \quad \frac{\partial v}{\partial y} = \frac{\frac{\partial f}{\partial u}}{\frac{\partial(f,g)}{\partial(u,v)}}$$

Quisiéramos hacer algunas observaciones sobre este resultado para darle unidad, usando para ello, en lugar de las funciones f y g, la función $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $F = (f, g)$.

Por principio nótese la matriz involucrada en el jacobiano $\frac{\partial(f,g)}{\partial(u,v)}$.

$$\frac{\partial(f,g)}{\partial(u,v)} = \det\begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix}$$

Esta matriz no es más que *la derivada de la función* $F = (f, g)$ (es decir, la matriz que representa a la transformación lineal $F'(u, v)$ en relación a la base canónica de $\mathbb{R}^2$ —ver sección 3 de este capítulo). Así entonces, que el jacobiano $\frac{\partial(f,g)}{\partial(u,v)}$ sea no nulo en $(\bar{u}, \bar{v})$, significa que la matriz jacobiana $JF(\bar{u}, \bar{v})$ es *inversible* (pues su determinante es distinto de cero).

Hagamos una observación similar con la función inversa $F^{-1}(x, y)$. Esta tiene por funciones coordenadas a las funciones $u = \varphi(x, y)$, $v = \psi(x, y)$. Es decir $F^{-1}(x, y) = (u, v) = (\varphi(x, y), \psi(x, y))$. La matriz jacobiana de esta función es

$$JF^{-1} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$

El resultado que acabamos de obtener nos dice cómo calcular las derivadas parciales $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ en una vecindad V de $(\bar{x}, \bar{y})$. Sustituyamos las fórmulas correspondientes en JF^{-1}, recordando que $\frac{\partial(f,g)}{\partial(u,v)} = \det(JF)$

$$JF^{-1} = \begin{bmatrix} \dfrac{\frac{\partial g}{\partial v}}{\det(JF)} & \dfrac{-\frac{\partial g}{\partial u}}{\det(JF)} \\ \dfrac{-\frac{\partial f}{\partial v}}{\det(JF)} & \dfrac{\frac{\partial f}{\partial u}}{\det(JF)} \end{bmatrix} = \frac{1}{\det(JF)}\begin{bmatrix} \frac{\partial g}{\partial v} & -\frac{\partial f}{\partial v} \\ -\frac{\partial g}{\partial u} & \frac{\partial f}{\partial u} \end{bmatrix}$$

Multipliquemos JF y JF^{-1}. Se obtiene

$$\begin{aligned}
(JF)(JF^{-1}) &= \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix}\left(\frac{1}{\det(JF)}\right)\begin{bmatrix} \frac{\partial g}{\partial v} & -\frac{\partial f}{\partial v} \\ -\frac{\partial g}{\partial u} & \frac{\partial f}{\partial u} \end{bmatrix} \\
&= \frac{1}{\det(JF)}\begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix}\begin{bmatrix} \frac{\partial g}{\partial v} & -\frac{\partial f}{\partial v} \\ -\frac{\partial g}{\partial u} & \frac{\partial f}{\partial u} \end{bmatrix} \\
&= \frac{1}{\det(JF)}\begin{bmatrix} \frac{\partial f}{\partial u}\frac{\partial g}{\partial v} - \frac{\partial f}{\partial v}\frac{\partial g}{\partial u} & 0 \\ 0 & -\frac{\partial g}{\partial u}\frac{\partial f}{\partial v} + \frac{\partial g}{\partial v}\frac{\partial f}{\partial u} \end{bmatrix} \\
&= \frac{\frac{\partial f}{\partial u}\frac{\partial g}{\partial v} - \frac{\partial f}{\partial v}\frac{\partial g}{\partial u}}{\det(JF)}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

Así, siendo el producto de la matriz JF por la matriz JF^{-1} igual a la matriz identidad, concluimos que la matriz jacobiana de la función inversa de F es justamente la inversa de la matriz jacobiana

de F (esto se puede concluir directamente al ver la expresión obtenida para la matriz JF^{-1} y recordando que si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es una matriz inversible, entonces su inversa se puede encontrar como $A^{-1} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$). Es decir, se tiene

$$JF^{-1} = (JF)^{-1}$$

De este modo, el resultado que conocíamos de nuestro primer curso de cálculo sigue siendo cierto en este contexto más general: la derivada de la función inversa F^{-1} es la inversa (en el sentido algebraico —inversa de una matriz) de la derivada de la función original

Recapitulando: si $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ es una función tal que $F(\bar{u}, \bar{v}) = (\bar{x}, \bar{y})$ y en una bola B de $(\bar{u}, \bar{v})$ las derivadas parciales $\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}$ de las funciones coordenadas de F son continuas, se tiene que siendo det $JF(\bar{u}, \bar{v}) \neq 0$, entonces existe una bola B' de $(\bar{x}, \bar{y})$ en la que existe la inversa F^{-1} de la función F, la cual tiene continuas las derivadas parciales de sus funciones coordenadas en B' y su matriz jacobiana es

$$JF^{-1}(x, y) = (JF(u, v))^{-1}$$

donde $(x, y) = (f(u, v), g(u, v)) \in B'$.

NOTA: Recordemos que si las funciones $x = f(u, v)$, $y = g(u, v)$ tienen derivadas parciales continuas en una bola B, decimos que son de clase $\mathscr{C}^1$. Siendo éstas las funciones coordenadas de F, podemos decir que esta función es de clase $\mathscr{C}^1$. Así, las hipótesis del resultado anterior se pueden establecer como: "si existe una bola B de centro en $(\bar{u}, \bar{v})$ en la que F es de clase $\mathscr{C}^1$ y si la derivada F' en $(\bar{u}, \bar{v})$ es inversible. . .". Estas mismas observaciones se hacen para la función F^{-1}. De esta manera el resultado anterior se puede establecer como: si $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ es tal que $F(\bar{u}, \bar{v}) = (\bar{x}, \bar{y})$ y en una bola B de $(\bar{u}, \bar{v})$, la función F es de clase $\mathscr{C}^1$, entonces si det $JF(\bar{u}, \bar{v}) \neq 0$, existe una bola B' de $(\bar{x}, \bar{y})$ en la cual existe la inversa F^{-1}, que es de clase $\mathscr{C}^1$ en B' y cuya derivada es $JF^{-1}(x, y) = (JF(u, v))^{-1}$ en donde $(x, y) = (f(u, v), g(u, v)) \in B'$.

Veamos algunos ejemplos.

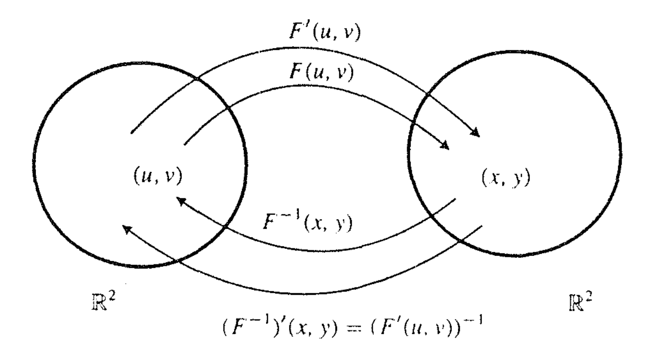


Figura 5. La función su inversa y sus derivadas.

Ejemplo 1. Considere la función $F:\mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(u, v) = (u^3 + v^3, u^2 + uv)$. Se tiene $F(1, 2) = (9, 3)$. Esta función es de clase $\mathscr{C}^1$ en $\mathbb{R}^2$. Las derivadas parciales de sus funciones coordenadas $x = f(u, v) = u^3 + v^3$ y $y = g(u, v) = u^2 + uv$ son

$$\frac{\partial f}{\partial u} = 3u^2, \quad \frac{\partial f}{\partial v} = 3v^2, \quad \frac{\partial g}{\partial u} = 2u + v, \quad \frac{\partial g}{\partial v} = u$$

La matriz jacobiana de F es

$$JF = \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix} = \begin{bmatrix} 3u^2 & 3v^2 \\ 2u + v & u \end{bmatrix}$$

la cual en el punto (1,2) es inversible pues

$$\det JF(1, 2) = \det \begin{bmatrix} 3 & 12 \\ 4 & 1 \end{bmatrix} = -45 \neq 0$$

Así, podemos concluir que en una bola B' de (9, 3) se da la inversa F^{-1} de F (o bien, que podemos despejar de $x = u^3 + v^3$, $y = u^2 + uv$, a u y v como funciones de x y y), la cual es de clase $\mathscr{C}^1$ en B', y que su derivada es

$$\begin{aligned} JF^{-1}(x, y) = (JF(u, v))^{-1} &= \frac{1}{\det JF} \begin{bmatrix} \frac{\partial g}{\partial v} & -\frac{\partial f}{\partial v} \\ -\frac{\partial g}{\partial u} & \frac{\partial f}{\partial u} \end{bmatrix} \\ &= \frac{1}{3u^3 - 6uv^2 - 3v^3} \begin{bmatrix} u & -3v^2 \\ -(2u + v) & 3u^2 \end{bmatrix} \end{aligned}$$

donde $x = u^3 + v^3$, $y = u^2 + uv$. Es decir

$$\begin{aligned} \frac{\partial u}{\partial x}(u^3 + v^3, u^2 + uv) &= \frac{u}{3u^3 - 6uv^2 - 3v^3} \\ \frac{\partial u}{\partial y}(u^3 + v^3, u^2 + uv) &= \frac{-3v^2}{3u^3 - 6uv^2 - 3v^3} \\ \frac{\partial v}{\partial x}(u^3 + v^3, u^2 + uv) &= -\frac{2u + v}{3u^3 - 6uv^2 - 3v^3} \\ \frac{\partial v}{\partial y}(u^3 + v^3, u^2 + uv) &= \frac{3u^2}{3u^3 - 6uv^2 - 3v^3} \end{aligned}$$

■

Ejemplo 2. Considere las ecuaciones $x = e^{u+v}$, $y = e^{u-v}$ que definen a x y y como funciones de u y v. Se quiere estudiar, a la luz de los resultados estudiados en esta sección, la posibilidad de despejar de ellas a u y v en términos de x y y. La función $F:\mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(u, v) = (e^{u+v}, e^{u-v})$ es diferenciable siempre y su derivada es

$$JF(u, v) = \begin{bmatrix} e^{u+v} & e^{u+v} \\ e^{u-v} & -e^{u-v} \end{bmatrix}$$

Como

$$\det JF(u, v) = -2e^{2u} \neq 0$$

concluimos que *siempre* (en cualquier punto) es posible despejar a u y v en términos de x y y. Más aún, como

$$(JF(u, v))^{-1} = \begin{bmatrix} \dfrac{e^{u-v}}{2e^{2u}} & \dfrac{e^{u+v}}{2e^{2u}} \\ \dfrac{e^{u-v}}{2e^{2u}} & -\dfrac{e^{u+v}}{2e^{2u}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}e^{-u-v} & \frac{1}{2}e^{-u+v} \\ \frac{1}{2}e^{-u-v} & -\frac{1}{2}e^{-u+v} \end{bmatrix}$$

concluimos que las derivadas parciales de las funciones inversas $u = \varphi(x, y)$, $v = \psi(x, y)$ son (en el punto (e^{u+v}, e^{u-v}))

$$\frac{\partial u}{\partial x} = \frac{1}{2}e^{-u-v}, \quad \frac{\partial u}{\partial y} = \frac{1}{2}e^{-u+v}, \quad \frac{\partial v}{\partial x} = \frac{1}{2}e^{-u-v}, \quad \frac{\partial v}{\partial y} = -\frac{1}{2}e^{-u+v}$$

Observe que en este caso es posible hacer explícitas las funciones $u = \varphi(x, y)$, $v = \psi(x, y)$. En efecto, de las expresiones $x = e^{u+v}$, $y = e^{u-v}$, se deduce que $u = \frac{1}{2}(\ln x + \ln y)$, $v = \frac{1}{2}(\ln x - \ln y)$. Las derivadas parciales de estas funciones son

$$\frac{\partial u}{\partial x} = \frac{1}{2}\frac{1}{x}, \quad \frac{\partial u}{\partial y} = \frac{1}{2}\frac{1}{y}, \quad \frac{\partial v}{\partial x} = \frac{1}{2}\frac{1}{x}, \quad \frac{\partial v}{\partial y} = -\frac{1}{2}\frac{1}{y}$$

las cuales coinciden con las obtenidas de la matriz jacobiana $(JF(u, v))^{-1}$ poniendo $x = e^{u+v}$, $y = e^{u-v}$. ■

Enunciemos ahora el resultado correspondiente para funciones $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ que generaliza las discusiones aquí presentadas en el caso $n = 2$.

Teorema 3.6.1 (De la Función Inversa). Sea $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $F(\mathbf{p}) = \mathbf{q}$, $\mathbf{p} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$, $\mathbf{q} = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$ Suponga que en una bola B de $\mathbb{R}^n$ con centro en $\mathbf{p}$ la función F es de clase $\mathscr{C}^1$ y que $\det JF(\mathbf{p}) \neq 0$. Entonces hay una bola B' en $\mathbb{R}^n$ con centro en $\mathbf{q}$ en la que se puede definir la función inversa de F, $F^{-1}: B' \to B$, la cual es de clase $\mathscr{C}^1$ y

$$JF^{-1}(y) = (JF(x))^{-1}$$

donde $y = F(x) \in B'$. ■

Ejemplo 3. Considere las ecuaciones

$$x = u + v + e^{w}$$
$$y = u + w + e^{2v}$$
$$z = v + w + e^{3u}$$

Para $\mathbf{p} = (\bar{u}, \bar{v}, \bar{w}) = (0, 0, 0)$ se tiene $\mathbf{q} = (\bar{x}, \bar{y}, \bar{z}) = (1, 1, 1)$. El determinante de la matriz jacobiana de la función $F(u, v, w) = (x, y, z)$ es

$$\det JF = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \det \begin{bmatrix} 1 & 1 & e^{w} \\ 1 & 2e^{2v} & 1 \\ 3e^{3u} & 1 & 1 \end{bmatrix}$$

que en el punto **p** es

$$\det JF(\mathbf{p}) = \det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} = -2 \neq 0$$

Entonces podemos, localmente, invertir la función F, en torno al punto $\mathbf{q}$, donde podemos definir funciones de clase $\mathscr{C}^1$ $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$. Como

$$JF^{-1}(q) = (JF(q))^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1/2 & 0 & 1/2 \\ -1 & 1 & 0 \\ 5/2 & -1 & -1/2 \end{bmatrix}$$

concluimos que las derivadas parciales de las funciones u, v, w en el punto $\mathbf{q} = (1, 1, 1)$ son

$$\begin{array}{lll} \dfrac{\partial u}{\partial x}(\mathbf{q}) = -\dfrac{1}{2}, & \dfrac{\partial u}{\partial y}(\mathbf{q}) = 0, & \dfrac{\partial u}{\partial z}(\mathbf{q}) = \dfrac{1}{2}, \\ \dfrac{\partial v}{\partial x}(\mathbf{q}) = -1, & \dfrac{\partial v}{\partial y}(\mathbf{q}) = 1, & \dfrac{\partial v}{\partial z}(\mathbf{q}) = 0, \\ \dfrac{\partial w}{\partial x}(\mathbf{q}) = \dfrac{5}{2}, & \dfrac{\partial w}{\partial y}(\mathbf{q}) = -1, & \dfrac{\partial w}{\partial z}(\mathbf{q}) = -\dfrac{1}{2} \end{array}$$ ■

Ejercicios (Capítulo 3, Sección 6)

En los ejercicios 1–5, constate que las funciones dadas $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ tienen inversa en los alrededores del punto $\mathbf{p} \in U$ dado. En cada caso determine la matriz jacobiana $JF^{-1}(F(\mathbf{p}))$.

1. $F(x, y) = (x + y, x - y)$, $\mathbf{p} = (x_0, y_0)$

2. $F(x, y) = (y, x)$, $\mathbf{p} = (x_0, y_0)$

3. $F(x, y) = (x^y, y^x)$, $\mathbf{p} = (1, 1)$

4. $F(x, y) = (x \operatorname{sen} y, y \cos x)$, $\mathbf{p} = (1, 1)$

5. $F(x, y) = (x + \arctan y, y + \arctan x)$, $\mathbf{p} = (1, 1)$

6. Considere la función $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (ax, by)$, donde a y b son reales positivos dados. Demuestre que esta función tiene inversa $F^{-1}: \mathbb{R}^2 \to \mathbb{R}^2$. Hállela. Determine los jacobianos $\det JF(x, y)$, $\det JF^{-1}(ax, by)$.

7. Repita el ejercicio anterior con la función $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (e^x, e^y)$.

8. Los dos ejercicios anteriores son casos particulares del siguiente resultado general: sean f, $g: \mathbb{R} \to \mathbb{R}$ funciones de clase $\mathscr{C}^1$ con derivada siempre positiva. La función $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (f(x), g(y))$ tiene inversa $F^{-1}: \mathbb{R}^2 \to \mathbb{R}^2$, la cual es $F^{-1}(u, v) = (f^{-1}(u), g^{-1}(v))$, donde $u = f(x)$, $v = g(y)$. Demuestre este hecho.

9. Sean $f, g: \mathbb{R} \to \mathbb{R}$ dos funciones de clase $\mathscr{C}^1$. Suponga que f no tiene puntos críticos y que g no tiene raíces. Demuestre que la función $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (f(x)g(y), f(y))$ tiene inversa. Determine el jacobiano $\det JF^{-1}(u, v)$.

10. Considere la función $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (x + y, xy)$. Demuestre que en los alrededores del punto (2, 1) es posible definir una inversa de F, $F^{-1}(u, v) = (x(u, v), y(u, v))$. Determine la matriz jacobiana $JF^{-1}(3, 2)$. Haga explícitas las funciones $x = x(u, v)$, $y = y(u, v)$ y verifique el resultado de la matriz $JF^{-1}(3, 2)$ derivando directamente estas funciones.

11. Considere la función de clase $\mathscr{C}^1$, $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (u, v) = (u(x, y), v(x, y))$. Suponga que las funciones $u, v: \mathbb{R}^2 \to \mathbb{R}$ tienen por gradientes en el origen a $\operatorname{grad} u(0, 0) = (3, 1)$, $\operatorname{grad} v(0, 0) = (-1, 2)$. Si $F(0, 0) = (0, 0)$, demuestre que existen funciones de clase $\mathscr{C}^1$ $x, y: B \subseteq \mathbb{R}^2 \to \mathbb{R}$, $x = x(u, v)$, $y = y(u, v)$ definidas en una bola B del origen, de modo que $F(x(u, v), y(u, v)) = (u, v)\ \forall (u, v) \in B$. Calcule $\operatorname{grad} x(0, 0)$, $\operatorname{grad} y(0, 0)$.

12. Sea $g: \mathbb{R} \to \mathbb{R}$ una función continua tal que $g(0) = 1$. Considere la función $F: \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$F(x, y) = \left(\int_x^y g(t)dt, \int_y^{x^2} g(t)dt \right)$$

Demuestre que esta función tiene una inversa F^{-1} definida en una bola B del origen de coordenadas. Determine $JF^{-1}(0, 0)$.

13. Sea $F: \mathbb{R}^3 \to \mathbb{R}^3$ la función $F(x, y, z) = (x + y + z, y + z, z)$. Demuestre que esta función tiene una inversa $F^{-1}: \mathbb{R}^3 \to \mathbb{R}^3$. Hállela.

14. Sean $f, g: \mathbb{R} \to \mathbb{R}$ dos funciones de clase $\mathscr{C}^1$. Considere la función $F: \mathbb{R}^3 \to \mathbb{R}^3$, $F(x, y, z) = (x + f(y) + g(z), y + g(z), z)$. Demuestre que F tiene inversa $F^{-1}: \mathbb{R}^3 \to \mathbb{R}^3$. Hállela. Determine la matriz jacobiana $JF^{-1}(u, v, w)$, donde $(u, v, w) = F(x, y, z)$.

(*) **15.** Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ definida en el abierto U de $\mathbb{R}^n$. En este ejercicio se probará que f *no puede ser inyectiva*. Sea $\mathbf{p} \in U$ un punto en el que $\operatorname{grad} f(\mathbf{p}) \neq 0$. (Si este punto no existe, ya no hay nada que demostrar. ¿Por qué?). Digamos que $\frac{\partial f}{\partial x_1}(\mathbf{p}) \neq 0$. Considere la función $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $F(x) = F(x_1, x_2, \ldots, x_n) = (f(x), x_2, x_3, \ldots, x_n)$. Demuestre que esta función tiene una inversa F^{-1} definida en alguna bola B de $F(\mathbf{p}) \in \mathbb{R}^n$. Concluya de aquí que existe una infinidad de puntos $\mathbf{x} \in \mathbb{R}^n$ tales que $f(\mathbf{x}) = c$, $c \in$ rango de f. Es decir, concluya que f no es inyectiva.

16. Demuestre que la función identidad $F: \mathbb{R}^n \to \mathbb{R}^n$, $F(\mathbf{x}) = \mathbf{x}$ es inversible. Determine su inversa $F^{-1}: \mathbb{R}^n \to \mathbb{R}^n$. Escriba también las matrices jacobianas $JF(\mathbf{x})$, $JF^{-1}(\mathbf{x})$.

17. Sea $F: \mathbb{R}^n \to \mathbb{R}^n$ la función $F(x_1, x_2, \ldots, x_n) = (x_n, x_{n-1}, \ldots, x_1)$. Demuestre que F tiene inversa $F^{-1}: \mathbb{R}^n \to \mathbb{R}^n$. Hállela. Escriba también las matrices jacobianas $JF(\mathbf{x})$, $JF^{-1}(\mathbf{x})$.

18. Sea $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una función de clase $\mathscr{C}^1$, definida en el abierto U de $\mathbb{R}^n$,

$$F(u_1, u_2, \ldots, u_n) = (x_1, x_2, \ldots, x_n) = \\ \big(x_1(u_1, u_2, \ldots, u_n), x_2(u_1, u_2, \ldots, u_n), \ldots, x_n(u_1, u_2, \ldots, u_n)\big)$$

tal que el determinante jacobiano $\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)}$ es distinto de cero en todo U. Considere la función inversa

$$F^{-1}(x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_n) = \\ \big(u_1(x_1, x_2, \ldots, x_n), u_2(x_1, x_2, \ldots, x_n), \ldots, u_n(x_1, x_2, \ldots, x_n)\big)$$

Demuestre que

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} \frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} = 1$$

(*) 3.7 Un interludio numérico: El método de Newton para sistemas no lineales

En esta sección entraremos en contacto con el mundo de los cálculos numéricos relacionados con el aparentemente simple problema de "resolver ecuaciones". Los antecedente escolares de este problema se pueden remontar a nuestros estudios del primer curso de álgebra (en la secundaria): se quiere obtener el valor de x que, por ejemplo, satisfaga (sea solución de) la ecuación $3x - 2 = 0$. Desde ese entonces se nos empezó a dar herramientas para resolver este tipo de problemas (en el caso mencionado: "pase el 2 sumando al miembro derecho y el 3 que multiplica a la x dividiendo, obteniendo así $x = 2/3$). Esta historia, cuyo comienzo ahora se ve "ingenuo", llega a tomar cauces en los que se presentan problemas altamente complicados. Ya con las ecuaciones algebraicas, o, con más precisión, ecuaciones del tipo $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$, sabemos que el gusto de tener de manera explícita las soluciones de ellas con una fórmula (que involucra radicales) se termina, en general, con $n = 4$. Uno de los grandes legados del matemático francés Evaristo Galois (1811–1832) fue precisamente este resultado (para $n \geq 5$ las ecuaciones mencionadas no tienen soluciones explícitas en términos de radicales). Este es un bellísimo e importante resultado teórico que, por otra parte, nos pone en un problema cuando nos enfrentamos con ecuaciones como $3x^7 - 2x^5 - 4x^2 + x + 2 = 0$. Es el "Análisis Numérico" quien se encarga de proporcionar métodos adecuados para obtener, no con una fórmula, pero sí con una aproximación numérica, las soluciones de la ecuación. El problema que abordaremos en esta sección está en la misma línea de resolver ecuaciones. Se trata de resolver *sistemas de ecuaciones*, es decir, conjuntos de n ecuaciones, cada una de las cuales involucra n incógnitas, y se trata de encontrar valores de éstas últimas que satisfagan todas y cada una de las ecuaciones dadas del sistema. Cuando estas ecuaciones son *lineales*, es decir, del tipo $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, el problema está bien estudiado por el Álgebra Lineal (es decir, se sabe en qué condiciones el sistema tiene solución y se conocen métodos para calcular tal solución). Si las ecuaciones del sistema no son lineales, el problema se vuelve mucho más severo. Este es el que queremos abordar en esta sección, pues con las ideas del cálculo en $\mathbb{R}^n$ que hemos desarrollado hasta este momento, es posible *entender* algunos de los métodos que ofrece el Análisis Numérico para resolver este tipo de problemas. En particular, estudiaremos un método numérico (el método de Newton), el cual es uno de los más eficaces para obtener soluciones de sistemas de ecuaciones no lineales.

El método de Newton (o de Newton-Raphson) para obtener raíces de ecuaciones del tipo $f(x) = 0$, con f derivable, se basa en lo siguiente. Se quiere encontrar el punto $\bar{x} \in \mathbb{R}$ donde $f(\bar{x}) = 0$. Partiendo de un punto x_1, como se muestra en la figura 1, trazamos la recta tangente a la curva $y = f(x)$ en el punto $(x_1, f(x_1))$, la cual es $y - f(x_1) = f'(x_1)(x - x_1)$. Esta recta corta el eje de las x en el punto $x_2 = x_1 - \frac{1}{f'(x_1)} f(x_1)$, el cual, en principio, se encuentra más cerca del punto $\bar{x}$ procurado[1]. Trazamos ahora la recta tangente a la curva $y = f(x)$ en el punto $(x_2, f(x_2))$ y localizamos el punto x_3 en que esta recta corta el eje de las x. Se obtiene que $x_3 = x_2 - \frac{1}{f'(x_2)} f(x_2)$, el cual se supone

[1]El método de Newton tiene algunas condiciones generales que aseguran su buen funcionamiento. Este tipo de detalles se estudian en cursos de Análisis Numérico. En esta sección no se discutirá esto; simplemente decimos que, "si se corre con buena suerte", nuestro punto de arranque en el método nos conducirá eventualmente a la raíz deseada.

que es una mejor aproximación del punto $\bar{x}$ procurado. Siguiendo con este proceso, es posible llegar eventualmente (con ciertas condiciones sobre la función f y sobre el punto inicial x_1) a la raíz buscada. Se tiene entonces, en general, que la fórmula

$$x_{n+1} = x_n - \frac{1}{f'(x_n)} f(x_n)$$

nos va dando los términos de una sucesión $x_1, x_2, \ldots$ la cual converge a la raíz $\bar{x}$.

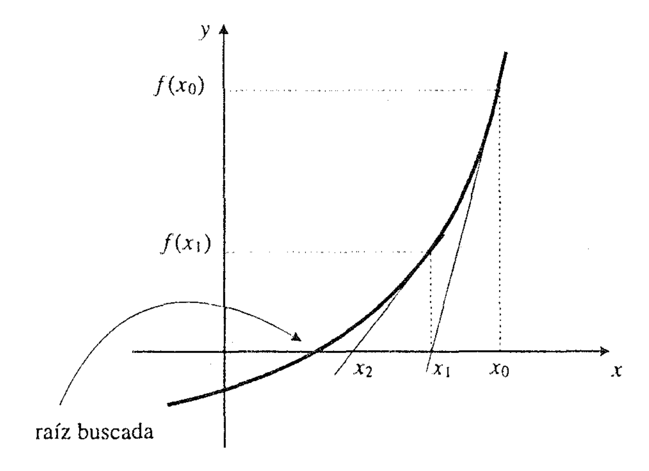


Figura 1. El Método de Newton para hallar raíces de ecuaciones $f(x) = 0$.

Veamos un ejemplo.

Ejemplo 1. Se quiere hallar una raíz de $f(x) = x - \ln(3x^2) - 3 = 0$. La derivada de esta función es $f'(x) = \frac{x-2}{x}$. En este caso la fórmula que nos dará las aproximaciones a la raíz es

$$x_{n+1} = x_n - \frac{x_n}{x_n - 2}(x_n - \ln(3x_n^2) - 3)$$

Partiendo del punto $x_1 = 4$, obtenemos los siguientes valores: $x_2 = 9.7424$, $x_3 = 8.36981$, $x_4 = 8.34098$, $x_5 = 8.34097 = x_6$. Así pues, con una exactitud de 5 cifras decimales, concluímos que la raíz buscada es $\bar{x} = 8.34097$. ■

Veamos con detenimiento la filosofía subyacente en el método de Newton anteriormente descrito. Tenemos entonces una curva en el plano, gráfica de la función $y = f(x)$, y estamos interesados en obtener el punto por donde esta curva corta el eje x. Partimos de un punto $\mathbf{p}_1 = (x_1, f(x_1))$ de la curva. *Si tal curva pudiera verse en el punto* $\mathbf{p}_1$ *como una recta, ésta debería ser la recta tangente a la curva en* $\mathbf{p}_1$. Lo que hacemos entonces es "ver" la curva $y = f(x)$ en $\mathbf{p}_1$, como una *función lineal,* la cual queda descrita por la ecuación de su recta tangente, y localizar el punto donde esta recta corta al eje x (este es un cálculo muy simple). La recta tangente a $y = f(x)$ en $\mathbf{p}_1$ es

$$y - f(x_1) = f'(x_1)(x - x_1)$$

que corta el eje x en $x_2 = x_1 - \frac{1}{f'(x_1)} f(x_1)$. Localizamos ahora el punto de la gráfica de la función $y = f(x)$ correspondiente a esta nueva abscisa. Este es $\mathbf{p}_2 = (x_2, f(x_2))$. Y repetimos la misma idea: aproximamos el comportamiento de la función $y = f(x)$ por un comportamiento lineal, descrito por su recta tangente en $\mathbf{p}_2$, la cual es

$$y - f(x_2) = f'(x_2)(x - x_2)$$

y vemos dónde corta ésta el eje x. Este punto es $x_3 = x_2 - \frac{1}{f'(x_2)} f(x_2)$. Y así sucesivamente. La idea que hay detrás de este proceso es entonces *aproximar el comportamiento de la función $y = f(x)$ por un comportamiento lineal y sobre este último ir localizando las raíces, las cuales se supone que se irán acercando a la raíz procurada de la función dada.*

Queremos abordar ahora un problema más general: encontrar una raíz de una función (bien portada desde el punto de vista de la diferenciabilidad) del tipo $F: \mathbb{R}^2 \to \mathbb{R}^2$, es decir, un punto $(\bar{x}, \bar{y})$ tal que $F(\bar{x}, \bar{y}) = (0, 0)$. Del mismo modo, si escribimos $F(x, y) = (f(x, y), g(x, y))$, con f, $g: \mathbb{R}^2 \to \mathbb{R}$ (las funciones coordenadas de F), se trata de encontrar un punto $(\bar{x}, \bar{y})$ donde $f(\bar{x}, \bar{y}) = 0$, $g(\bar{x}, \bar{y}) = 0$. Incluso, de otra manera: se trata de encontrar la solución $(\bar{x}, \bar{y})$ del sistema de ecuaciones dado por

$$f(x, y) = 0, \quad g(x, y) = 0$$

Si las funciones f y g fueran lineales, el problema sería muy simple de resolver. En efecto, digamos que

$$f(x, y) = ax + by + \alpha, \quad g(x, y) = cx + dy + \beta$$

Se trataría entonces de resolver el sistema de dos ecuaciones lineales con dos indeterminadas

$$ax + by + \alpha = 0, \quad cx + dy + \beta = 0$$

el cual está perfectamente estudiado desde los cursos elementales de álgebra en la secundaria. El caso de interés, que consideraremos aquí, es cuando f y g no son lineales. Para ello, usamos las mismas ideas del método de Newton que acabamos de recordar: partiendo de un punto $\mathbf{p}_1 = (x_1, y_1)$ —una primera aproximación a la raíz que buscamos—, consideramos que el comportamiento de f y g es lineal. Es decir, las funciones f y g se verían (en los alrededores de $\mathbf{p}_1$) como

$$f(x, y) = a(x - x_1) + b(y - y_1) + \alpha$$
$$g(x, y) = c(x - x_1) + d(y - y_1) + \beta$$

para ciertos números reales $a, b, c, d, \alpha, \beta$. Si en la primera de estas expresiones ponemos $x = x_1$, $y = y_1$, obtenemos que α debe ser $f(x_1, y_1)$. Igualmente, si derivamos esta expresión respecto de x y sustituimos x por x_1, y por y_1, obtenemos que a debe ser $\frac{\partial f}{\partial x}(\mathbf{p}_1)$. Así mismo, se obtiene que b debe ser $\frac{\partial f}{\partial y}(\mathbf{p}_1)$. De modo que, si quisiéramos ver a $f(x, y)$ como una función lineal en los alrededores del punto $\mathbf{p}_1 = (x_1, y_1)$, ésta debería ser

$$f(x, y) = f(\mathbf{p}_1) + \frac{\partial f}{\partial x}(\mathbf{p}_1)(x - x_1) + \frac{\partial f}{\partial y}(\mathbf{p}_1)(y - y_1)$$

De manera análoga, si quisiéramos ver a $g(x, y)$ como una función lineal en los alrededores del punto $\mathbf{p}_1$, ésta debe ser

$$g(x, y) = g(\mathbf{p}_1) + \frac{\partial g}{\partial x}(\mathbf{p}_1)(x - x_1) + \frac{\partial g}{\partial y}(\mathbf{p}_1)(y - y_1)$$

(observe que lo que estamos diciendo es que si los comportamientos de las superficies $z = f(x, y)$ y $z = g(x, y)$ en los alrededores del punto $\mathbf{p}_1$ fueran lineales, éstos serían los de sus planos tangentes en el punto $\mathbf{p}_1$). En términos de la función $F(x, y) = (f(x, y), g(x, y))$, si en los alrededores del punto $\mathbf{p}_1$ esta función tuviera un comportamiento lineal, éste sería dado por

$$\begin{aligned} F(x, y) &= \big(f(x, y), g(x, y)\big) \\ &= \left(f(\mathbf{p}_1) + \frac{\partial f}{\partial x}(\mathbf{p}_1)(x - x_1) + \frac{\partial f}{\partial y}(\mathbf{p}_1)(y - y_1), g(\mathbf{p}_1) + \frac{\partial g}{\partial x}(\mathbf{p}_1)(x - x_1)\right. \\ &\quad \left. + \frac{\partial g}{\partial y}(\mathbf{p}_1)(y - y_1)\right) \\ &= \big(f(\mathbf{p}_1), g(\mathbf{p}_1)\big) + \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}_1) & \frac{\partial f}{\partial y}(\mathbf{p}_1) \\ \frac{\partial g}{\partial x}(\mathbf{p}_1) & \frac{\partial g}{\partial y}(\mathbf{p}_1) \end{bmatrix} \begin{bmatrix} x - x_1 \\ y - y_1 \end{bmatrix} \\ &= F(\mathbf{p}_1) + JF(\mathbf{p}_1)\big((x, y) - \mathbf{p}_1\big) \end{aligned}$$

donde $JF(\mathbf{p}_1)$ es la matriz jacobiana de F en el punto $\mathbf{p}_1$, y estamos identificando al vector (a, b) con la matriz $\begin{bmatrix} a \\ b \end{bmatrix}$. Obsérvese la expresión obtenida para F en los alrededores de $\mathbf{p}_1$:

$$F(x, y) = F(\mathbf{p}_1) + JF(\mathbf{p}_1)\big((x, y) - \mathbf{p}_1\big)$$

y compare con la que habíamos establecido para la función $y = f(x)$ de nuestra discusión previa, en los alrededores del punto $\mathbf{p}_1$ de abscisa x_1, la cual es $y - f(x_1) = f'(x_1)(x - x_1)$, o bien

$$f(x) = f(x_1) + f'(x_1)(x - x_1)$$

Es claro que se trata de "la misma expresión" que establece la aproximación del comportamiento lineal de la función en los alrededores del punto $\mathbf{p}_1$: la derivada de la función F es, como ya habíamos dicho, representada por su matriz jacobiana JF.

Entonces, se tiene que la función F, si tuviera un comportamiento lineal en los alrededores del punto $\mathbf{p}_1$, estaría dado por

$$F(x, y) = F(\mathbf{p}_1) + JF(\mathbf{p}_1)\big((x, y) - \mathbf{p}_1\big)$$

De esta expresión hallamos la raíz $\mathbf{p}_2 = (x_2, y_2)$, la cual nos servirá como una nueva aproximación de la raíz que buscamos de F. Se tiene entonces que

$$F(\mathbf{p}_1) + JF(\mathbf{p}_1)\big((x_2, y_2) - \mathbf{p}_1\big) = 0$$

De aquí queremos despejar el punto (x_2, y_2). Nótese que en este caso no podemos "pasar dividiendo la derivada de F", como hacíamos en el caso del método de Newton para una variable, pues ahora la derivada es una matriz. Lo que sí podemos hacer, para quitarla de nuestro camino (y dejar solo el punto (x_2, y_2)), es tomar su inversa (suponiendo que ésta existiera). De hecho se tiene

$$\begin{aligned} JF(\mathbf{p}_1)\big((x_2, y_2) - \mathbf{p}_1\big) &= -F(\mathbf{p}_1) \\ \big((x_2, y_2) - \mathbf{p}_1\big) &= [JF(\mathbf{p}_1)]^{-1}(-F(\mathbf{p}_1)) = -[JF(\mathbf{p}_1)]^{-1}F(\mathbf{p}_1) \end{aligned}$$

de donde

$$(x_2, y_2) = (x_1, y_1) - [JF(x_1, y_1)]^{-1} F(x_1, y_1)$$

Siguiendo con este proceso (aproximándose ahora a la función F por un comportamiento lineal alrededor del punto $\mathbf{p}_2$ y hallando su raíz), llegamos finalmente a la fórmula del método de Newton (en dos dimensiones)

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) - [JF(x_n, y_n)]^{-1} F(x_n, y_n)$$

la cual nos da (en determinadas condiciones sobre F y del punto inicial $\mathbf{p}_1$) una sucesión de puntos $\mathbf{p}_n = (x_n, y_n)$ en $\mathbb{R}^2$ que se aproximan a la raíz $(\bar{x}, \bar{y})$ buscada de F. Obsérvese que en cada paso, para obtener un nuevo punto (x_{n+1}, y_{n+1}), es necesario calcular la matriz jacobiana de F en el punto anterior (x_n, y_n) e invertirla. Es posible, en este caso de dos dimensiones, dejar explícitas las cuentas que se encuentran involucradas en este proceso, con ayuda de la fórmula $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Se tiene

$$\begin{aligned}
[JF(x_n, y_n)]^{-1} &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}_n) & \frac{\partial f}{\partial y}(\mathbf{p}_n) \\ \frac{\partial g}{\partial x}(\mathbf{p}_n) & \frac{\partial g}{\partial y}(\mathbf{p}_n) \end{bmatrix}^{-1} \\
&= \frac{1}{\frac{\partial f}{\partial x}(\mathbf{p}_n)\frac{\partial g}{\partial y}(\mathbf{p}_n) - \frac{\partial f}{\partial y}(\mathbf{p}_n)\frac{\partial g}{\partial x}(\mathbf{p}_n)} \begin{bmatrix} \frac{\partial g}{\partial y}(\mathbf{p}_n) & -\frac{\partial f}{\partial y}(\mathbf{p}_n) \\ -\frac{\partial g}{\partial x}(\mathbf{p}_n) & \frac{\partial f}{\partial x}(\mathbf{p}_n) \end{bmatrix} \\
&= \frac{1}{\det JF(\mathbf{p}_n)} \begin{bmatrix} \frac{\partial g}{\partial y}(\mathbf{p}_n) & -\frac{\partial f}{\partial y}(\mathbf{p}_n) \\ -\frac{\partial g}{\partial x}(\mathbf{p}_n) & \frac{\partial f}{\partial x}(\mathbf{p}_n) \end{bmatrix}
\end{aligned}$$

de modo que

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \end{bmatrix} - \frac{1}{\det JF(\mathbf{p}_n)} \begin{bmatrix} \frac{\partial g}{\partial y}(\mathbf{p}_n) & -\frac{\partial f}{\partial y}(\mathbf{p}_n) \\ -\frac{\partial g}{\partial x}(\mathbf{p}_n) & \frac{\partial f}{\partial x}(\mathbf{p}_n) \end{bmatrix} \begin{bmatrix} f(\mathbf{p}_n) \\ g(\mathbf{p}_n) \end{bmatrix}$$

es decir

$$\begin{aligned}
x_{n+1} &= x_n - \frac{1}{\det JF(\mathbf{p}_n)}\left(f(\mathbf{p}_n)\frac{\partial g}{\partial y}(\mathbf{p}_n) - g(\mathbf{p}_n)\frac{\partial f}{\partial y}(\mathbf{p}_n)\right) \\
y_{n+1} &= y_n - \frac{1}{\det JF(\mathbf{p}_n)}\left(g(\mathbf{p}_n)\frac{\partial f}{\partial x}(\mathbf{p}_n) - f(\mathbf{p}_n)\frac{\partial g}{\partial x}(\mathbf{p}_n)\right)
\end{aligned}$$

Ejemplo 2. Se quiere resolver el sistema

$$\begin{aligned}
f(x, y) &= 5x^2 + 6xy + 5y^2 - 4x + 4y - 4 = 0 \\
g(x, y) &= x^2 + y^2 - 1 = 0
\end{aligned}$$

Desde el punto de vista geométrico, se tratan de hallar las intersecciones entre la elipse que representa $f(x, y) = 0$ (la cual, en un sistema coordenado con el eje x' siendo la recta $y = -x$, y el eje y' siendo

la recta $y = x - 2$, se ve como $(x' - 1)^2 + 4(y' + 1)^2 = 4$), y el círculo unitario $g(x, y) = 0$. La matriz jacobiana de la función $F(x, y) = (f(x, y), g(x, y))$ es

$$JF(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} = \begin{bmatrix} 10x + 6y - 4 & 6x + 10y + 4 \\ 2x & 2y \end{bmatrix}$$

cuyo determinante es

$$\det JF = 12y^2 - 12x^2 - 8x - 8y$$

Entonces, los puntos (x_n, y_n) de la sucesión que aproximará las raíces procuradas se calculan como

$$\begin{aligned} x_{n+1} &= x_n - \frac{1}{\det JF(\mathbf{p}_n)}\left(f(\mathbf{p}_n)\frac{\partial g}{\partial y}(\mathbf{p}_n) - g(\mathbf{p}_n)\frac{\partial f}{\partial y}(\mathbf{p}_n)\right) \\ &= x_n - \frac{1}{12y_n^2 - 12x_n^2 - 8x_n - 8y_n}\big((5x_n^2 + 6x_n y_n + 5y_n^2 - 4x_n + 4y_n - 4)(2y_n) \\ &\quad - (x_n^2 + y_n^2 - 1)(6x_n + 10y_n + 4)\big) \\ &= x_n - \frac{1}{12y_n^2 - 12x_n^2 - 8x_n - 8y_n}\left(6x_n y_n^2 + 4y_n^2 - 6x_n^3 - 8x_n y_n + 2y_n - 4x_n^2 + 6x_n + 4\right) \\ y_{n+1} &= y_n - \frac{1}{\det JF(\mathbf{p}_n)}\left(g(\mathbf{p}_n)\frac{\partial f}{\partial x}(\mathbf{p}_n) - f(\mathbf{p}_n)\frac{\partial g}{\partial x}(\mathbf{p}_n)\right) \\ &= y_n - \frac{1}{12y_n^2 - 12x_n^2 - 8x_n - 8y_n}\big((x_n^2 + y_n^2 - 1)(10x_n + 6y_n - 4) \\ &\quad - (5x_n^2 + 6x_n y_n + 5y_n^2 - 4x_n + 4y_n - 4)(2x_n)\big) \\ &= y_n - \frac{1}{12y_n^2 - 12x_n^2 - 8x_n - 8y_n}\left(-6x_n^2 y_n + 4x_n^2 + 6y_n^3 - 8x_n y_n - 2x_n - 4y_n^2 - 6y_n + 4\right) \end{aligned}$$

Partiendo del punto $\mathbf{p}_1 = (x_1, y_1) = (1, 1)$, se obtienen los siguientes puntos

$$\begin{aligned} \mathbf{p}_2 &= (1.25, 0.25) \\ \mathbf{p}_3 &= (0.991667, 0.291667) \\ \mathbf{p}_4 &= (0.957506, 0.290433) \\ \mathbf{p}_5 &= (0.956943, 0.290276) = \mathbf{p}_6 \end{aligned}$$

Así pues, una raíz del sistema corresponde al punto $(\tilde{x}, \tilde{y}) = (0.956943, 0.290276)$. Partiendo ahora del punto $\mathbf{p}_1 = (-1, -1)$ se obtienen los siguientes resultados

$$\begin{aligned} \mathbf{p}_2 &= (-0.25, -1.25) \\ \mathbf{p}_3 &= (-0.291667, -0.991667) \\ \mathbf{p}_4 &= (-0.290433, -0.957506) \\ \mathbf{p}_5 &= (-0.290276, -0.956943) = \mathbf{p}_6 \end{aligned}$$

Entonces otra raíz del sistema es $(\tilde{x}, \tilde{y}) = (-0.290276, -0.956943)$. ∎

Ejemplo 3. Se quiere resolver el sistema

$$f(x, y) = y^2 - (x - 1)^3 = 0$$
$$g(x, y) = \ln(1 + y^2) + e^{-x} - 3x + 2y = 0$$

La matriz jacobiana de la función $F: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (f(x, y), g(x, y))$ es

$$JF(x, y) = \begin{bmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{bmatrix} = \begin{bmatrix} -3(x-1)^2 & 2y \\ -e^{-x} - 3 & \dfrac{2y}{1+y^2} + 2 \end{bmatrix}$$

El determinante de esta matriz es

$$\det JF(x, y) = -\frac{6y(x-1)^2}{1+y^2} - 6(x-1)^2 + 2ye^{-x} + 6y$$

Entonces las fórmulas que dan las nuevas aproximaciones (x_n, y_n) a la raíz procurada son

$$x_{n+1} = x_n - \frac{1}{\det JF(x_n, y_n)}\left[\left(y_n^2 - (x_n - 1)^3\right)\left(\frac{2y_n}{1+y_n^2} + 2\right)\right.$$
$$\left. - 2y_n\left(\ln(1 + y_n^2) + \exp(-x_n) - 3x_n + 2y_n\right)\right]$$
$$y_{n+1} = y_n - \frac{1}{\det JF(x_n, y_n)}\left[-3(x_n - 1)^2\left(\ln(1 + y_n^2) + \exp(-x_n) - 3x_n + 2y_n\right)\right.$$
$$\left. + (\exp(-x_n) + 3)\left(y_n^2 - (x_n - 1)^3\right)\right]$$

Tomando como punto inicial a $(x_0, y_0) = (1, 1)$ se obtienen los siguientes puntos

$$\mathbf{p}_1 = (0.572736, 0.5)$$
$$\mathbf{p}_2 = (0.15438, -0.057117)$$
$$\mathbf{p}_3 = (0.417146, 0.330259)$$
$$\mathbf{p}_4 = (3.409549, 4.482523)$$
$$\mathbf{p}_5 = (3.309779, 3.607891)$$
$$\mathbf{p}_6 = (3.419417, 3.754896)$$
$$\mathbf{p}_7 = (3.412339, 3.746734)$$
$$\mathbf{p}_8 = (3.412304, 3.746693) = \mathbf{p}_9$$

Así, una raíz del sistema corresponde a los valores $x = 3.412304$, $y = 3.746693$. Es interesante notar que, aunque se parta de un punto muy alejado de la raíz, *en este caso* el método de Newton conduce rápidamente a las cercanías de la raíz procurada. Por ejemplo, partiendo del punto

$\mathbf{p}_0 = (x_0, y_0) = (30, 45)$, se obtienen los siguientes puntos en el proceso iterativo

$$\begin{aligned}
\mathbf{p}_1 &= (20.506144, 27.344458)\\
\mathbf{p}_2 &= (14.22879, 18.362055)\\
\mathbf{p}_3 &= (10.050146, 12.482694)\\
\mathbf{p}_4 &= (7.283054, 8.698204)\\
\mathbf{p}_5 &= (5.477638, 6.3161)\\
\mathbf{p}_6 &= (4.347231, 4.88236)\\
\mathbf{p}_7 &= (3.71613, 4.109406)\\
\mathbf{p}_8 &= (3.460018, 3.80308)\\
\mathbf{p}_9 &= (3.413773, 3.748422)\\
\mathbf{p}_{10} &= (3.412306, 3.746696)\\
\mathbf{p}_{11} &= (3.412304, 3.746693) = \mathbf{p}_{12}
\end{aligned}$$

■

Las ideas expuestas anteriormente para obtener la fórmula del método de Newton para funciones $F: \mathbb{R}^2 \to \mathbb{R}^2$, se extienden fácilmente para el caso de funciones $F: \mathbb{R}^n \to \mathbb{R}^n$, quedando de la siguiente manera: si $F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))$ —donde $\mathbf{x} = (x_1, x_2, \ldots, x_n)$— es una función (bien portada desde el punto de vista diferenciable), con funciones coordenadas $f_1, f_2, \ldots, f_n: \mathbb{R}^n \to \mathbb{R}$, para la cual se quiere hallar una raíz $\tilde{\mathbf{x}} \in \mathbb{R}^n$ (es decir, un vector $\tilde{\mathbf{x}} \in \mathbb{R}^n$ tal que $F(\tilde{\mathbf{x}}) = 0$), o, en forma equivalente, se quieren obtener valores $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n \in \mathbb{R}$ que satisfagan el sistema de n ecuaciones con n indeterminadas

$$\begin{aligned}
f_1(x_1, x_2, \ldots, x_n) &= 0\\
f_2(x_1, x_2, \ldots, x_n) &= 0\\
&\cdots\\
f_n(x_1, x_2, \ldots, x_n) &= 0
\end{aligned}$$

entonces la fórmula (del método de Newton)

$$\mathbf{p}_{m+1} = \mathbf{p}_m - (JF(\mathbf{p}_m))^{-1}F(\mathbf{p}_m)$$

da una sucesión de puntos $\mathbf{p}_1, \mathbf{p}_2, \ldots$ (partiendo de un punto $\mathbf{p}_0$ inicial) en $\mathbb{R}^n$ que tienden (en determinadas condiciones sobre F y sobre el punto de partida $\mathbf{p}_0$) a la raíz procurada.

En la fórmula anterior, $JF(\mathbf{p}_m)$ es la matriz (de orden $n \times n$) jacobiana de la función F evaludada en el punto $\mathbf{p}_m$. Es decir

$$JF(\mathbf{p}_m) = \begin{bmatrix}
\dfrac{\partial f_1}{\partial x_1}(\mathbf{p}_m) & \dfrac{\partial f_1}{\partial x_2}(\mathbf{p}_m) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{p}_m)\\
\dfrac{\partial f_2}{\partial x_1}(\mathbf{p}_m) & \dfrac{\partial f_2}{\partial x_2}(\mathbf{p}_m) & \cdots & \dfrac{\partial f_2}{\partial x_n}(\mathbf{p}_m)\\
\vdots & \vdots & & \vdots\\
\dfrac{\partial f_n}{\partial x_1}(\mathbf{p}_m) & \dfrac{\partial f_n}{\partial x_2}(\mathbf{p}_m) & \cdots & \dfrac{\partial f_n}{\partial x_n}(\mathbf{p}_m)
\end{bmatrix}$$

En el caso $n = 2$, hicimos explícita la fórmula de la inversa de esta matriz que está involucrada en la correspondiente fórmula del método de Newton. Es claro que en este caso general resulta completamente impráctico hacer explícita tal inversa (en términos de la adjunta y del determinante de la matriz), pues las expresiones que se obtendrían serían de manejo altamente complicado. En este caso general entonces, cada nueva iteración nos conduce, en principio, a invertir la matriz jacobiana con alguno de los métodos de eliminación estudiados en álgebra lineal.

Así pues, partiendo de un punto inicial $\mathbf{p}_0$ (que se supone debe estar "cerca" de la raíz procurada), la fórmula del método de Newton establece que para obtener el siguiente punto $\mathbf{p}_1$ en el proceso de aproximación a la raíz $\bar{\mathbf{x}}$ se deben hacer los siguientes cálculos:

1. Evaluar la matriz jacobiana JF en el punto $\mathbf{p}_0$.
2. Invertir la matriz obtenida en 1.
3. Multiplicar la matriz obtenida en 2 por el vector $F(\mathbf{p}_n)$ (es decir, por la matriz $n \times 1$ cuyos elementos son las coordenadas de $F(\mathbf{p}_n)$).
4. Hacer la resta $\mathbf{p}_0$ menos el vector obtenido en el paso anterior, llegando así al punto $\mathbf{p}_1$, con el cual se repite el proceso desde el paso 1.

El procedimiento anterior, si bien es claro de entender, involucra una gran cantidad de cálculos en cada paso (invertir matrices, multiplicar matrices, restar vectores). Un hecho que es interesante hacer notar, es que en tal proceso *no importa* el *resultado parcial del paso 2; lo que importa es el resultado del paso 3*, pues éste se usará en la etapa final 4. Es decir, no importa en sí tener la inversa $(JF(\mathbf{p}_m))^{-1}$, lo que importa es tener el producto $(JF(\mathbf{p}_m))^{-1}F(\mathbf{p}_m)$. Veamos cómo podemos sacar provecho de esta situación "accidental": llamemos $\mathbf{v}_m$ al vector $(JF(\mathbf{p}_m))^{-1}F(\mathbf{p}_m)$. *Este es el vector que interesa*. Tenemos entonces que

$$\mathbf{v}_m = (JF(\mathbf{p}_m))^{-1}F(\mathbf{p}_m)$$

o bien, multiplicando ambos miembros por $JF(\mathbf{p}_m)$, queda como

$$JF(\mathbf{p}_m)\mathbf{v}_m = F(\mathbf{p}_m)$$

Esta expresión se puede contemplar como un sistema no homogéneo de n ecuaciones lineales con n indeterminadas (las coordenadas del vector $\mathbf{v}_m$), siendo $JF(\mathbf{p}_m)$ la matriz de coeficientes del sistema y $F(\mathbf{p}_m)$ la matriz de términos independientes. Este sistema puede ser resuelto con alguno de los métodos de eliminación estudiados en el álgebra lineal. Aunque en apariencia llegamos a una situación similar a la presentada originalmente, en la que se tuvo que echar mano de procesos de eliminación Gaussiana (en la situación original para invertir una matriz, y en esta nueva situación para resolver un sistema de ecuaciones lineales), resulta que esta nueva manera de ver las cosas tiene, "en la práctica", grandes ventajas sobre la primera. La cuestión es que este proceso puede ser implementado en un programa por computadora, y aquí es donde la nueva visión de la fórmula del método de Newton adquiere ventajas. Más adelante entramos en detalles sobre la manera de estructurar un programa que haga todos los cálculos involucrados en el método de Newton.

En resumen, para obtener una raíz de $F(x) = 0$, se procede como sigue: partiendo de un punto $\mathbf{p}_0$, hacemos los siguientes pasos para $m = 0, 1, 2, \ldots$

1. Resolvemos el sistema $JF(\mathbf{p}_m)\mathbf{v}_m = F(\mathbf{p}_m)$ para el vector $\mathbf{v}_m$.
2. Hacemos $\mathbf{p}_{m+1} = \mathbf{p}_m - \mathbf{v}_m$.
3. Regresamos a 1) con el punto $\mathbf{p}_{m+1}$.

Ejemplo 4. Consideremos el sistema

$$f(x, y, z) = x^2 + y^2 + \tanh z - 3 = 0$$
$$g(x, y, z) = x + \tanh(2y) - z^3 - 2 = 0$$
$$h(x, y, z) = \tanh(3x) + y^3 + z - 1 = 0$$

La matriz jacobiana de la función $F: \mathbb{R}^3 \to \mathbb{R}^3$, $F(x, y, z) = (f(x, y, z), g(x, y, z), h(x, y, z))$ es

$$JF(x, y, z) = \begin{bmatrix} 2x & 2y & \operatorname{sech}^2 z \\ 1 & 2\operatorname{sech}^2 2y & -3z^2 \\ 3\operatorname{sech}^2 3x & 3y^2 & 1 \end{bmatrix}$$

Partiendo del punto $\mathbf{p}_0 = (1, 1, 1)$, en cada paso se obtiene el vector $\mathbf{v}_m = (\alpha_m, \beta_m, \gamma_m)$, solución del sistema

$$\begin{bmatrix} 2x_m & 2y_m & \operatorname{sech}^2 z_m \\ 1 & 2\operatorname{sech}^2 2y_m & -3z_m^2 \\ 3\operatorname{sech}^2 3x_m & 3y_m^2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_m \\ \beta_m \\ \gamma_m \end{bmatrix} \begin{bmatrix} x_m^2 + y_m^2 + \tanh z_m - 3 \\ x_m + \tanh 2y_m - z_m^3 - 2 \\ \tanh 3x_m + y_m^3 + z - 1 \end{bmatrix}$$

y calculando el siguiente punto $\mathbf{p}_{m+1}$ como $\mathbf{p}_m - \mathbf{v}_m$, $m = 0, 1, 2, \ldots$, llegando a los siguientes resultados

$$m = 1, \quad \mathbf{v}_0 = \begin{bmatrix} -0.777585 \\ 0.634025 \\ 0.115991 \end{bmatrix}, \quad \mathbf{p}_1 = \mathbf{p}_0 - \mathbf{v}_0 = \begin{bmatrix} 1.77758 \\ 0.365975 \\ 0.884009 \end{bmatrix}$$

$$m = 2, \quad \mathbf{v}_1 = \begin{bmatrix} 0.021336 \\ 0.867907 \\ 0.584239 \end{bmatrix}, \quad \mathbf{p}_2 = \mathbf{p}_1 - \mathbf{v}_1 = \begin{bmatrix} 1.75624 \\ -0.501932 \\ 0.29977 \end{bmatrix}$$

$$m = 3, \quad \mathbf{v}_2 = \begin{bmatrix} -0.222095 \\ -0.73653 \\ -0.730008 \end{bmatrix}, \quad \mathbf{p}_3 = \mathbf{p}_2 - \mathbf{v}_2 = \begin{bmatrix} 1.97834 \\ 0.234598 \\ -0.430237 \end{bmatrix}$$

$$m = 4, \quad \mathbf{v}_3 = \begin{bmatrix} 0.228798 \\ 0.020467 \\ -0.420724 \end{bmatrix}, \quad \mathbf{p}_4 = \mathbf{p}_3 - \mathbf{v}_3 = \begin{bmatrix} 1.74954 \\ 0.214131 \\ -0.009514 \end{bmatrix}$$

$$m = 5, \quad \mathbf{v}_4 = \begin{bmatrix} 0.0211528 \\ 0.0790144 \\ -0.010626 \end{bmatrix}, \quad \mathbf{p}_5 = \mathbf{p}_4 - \mathbf{v}_4 = \begin{bmatrix} 1.72839 \\ 0.135116 \\ 0.001113 \end{bmatrix}$$

$$m = 6, \quad \mathbf{v}_5 = \begin{bmatrix} 0.0012225 \\ -0.004832 \\ 0.0037808 \end{bmatrix}, \quad \mathbf{p}_6 = \mathbf{p}_5 - \mathbf{v}_5 = \begin{bmatrix} 1.72717 \\ 0.139949 \\ -0.002668 \end{bmatrix}$$

$$m = 7, \quad \mathbf{v}_6 = \begin{bmatrix} 1.08104 \times 10^{-5} \\ -1.62693 \times 10^{-5} \\ 9.51842 \times 10^{-7} \end{bmatrix}, \quad \mathbf{p}_7 = \mathbf{p}_6 - \mathbf{v}_6 = \begin{bmatrix} 1.72715 \\ 0.139965 \\ -0.00266914 \end{bmatrix}$$

La siguiente iteración arroja un punto que difiere de los elementos de $\mathbf{p}_7$ en orden de 10^{-6}. Así, una raíz del sistema es $x = 1.72715$, $y = 0.139965$, $z = -0.00266914$. ■

En la siguiente página se muestra el diagrama de flujo de un programa de computadora que hace los cálculos del método de Newton. La parte importante corresponde a un procedimiento en el que, en cada nueva iteración, se resuelve el sistema lineal de ecuaciones $JF(\mathbf{p}_m)\mathbf{v}_m = F(\mathbf{p}_m)$ para el vector $\mathbf{v}_m$, que servirá para calcular la siguiente aproximación x_{m+1}.

Ejemplo 5. Se quiere resolver el sistema

$$f_1(x, y, z, u) = x^3 + y^3 - 2z^2 + u^2 + 3 = 0$$
$$f_2(x, y, z, u) = xy - zu + 2 = 0$$
$$f_3(x, y, z, u) = x + y^2 + z + u = 0$$
$$f_4(x, y, z, u) = xz + yu - 2 = 0$$

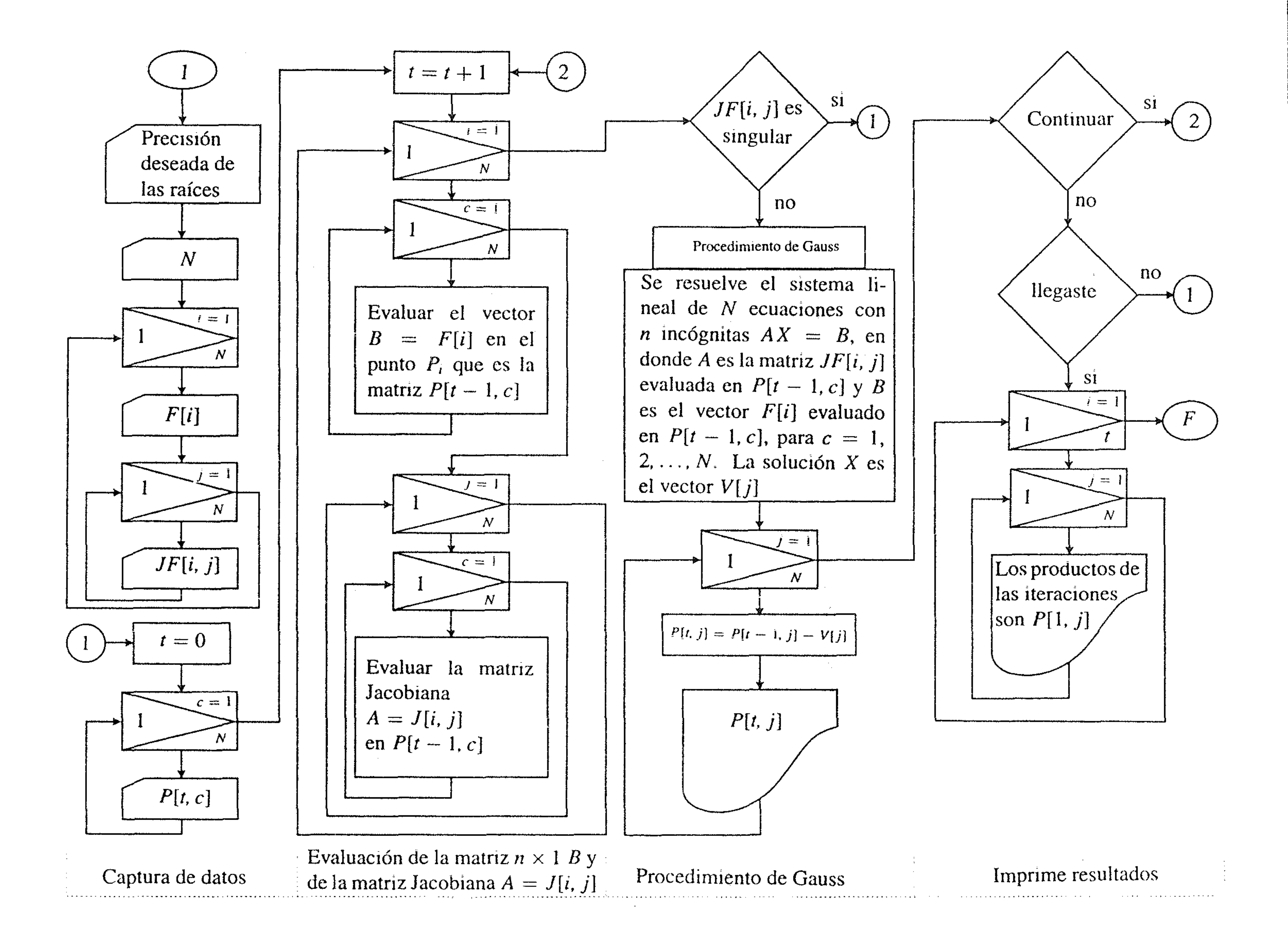
La matriz jacobiana de la función $F\colon \mathbb{R}^4 \to \mathbb{R}^4$, $F = (f_1, f_2, f_3, f_4)$ es

$$JF(x, y, z, u) = \begin{bmatrix} 3x^2 & 3y^2 & -4z & 2u \\ y & x & -u & -z \\ 1 & 2y & 1 & 1 \\ z & u & x & y \end{bmatrix}$$

A continuación se muestran los resultados de las iteraciones realizadas por el programa mencionado anteriormente, con el punto inicial $x = 1$, $y = 2$, $z = 3$, $u = 4$.

iteración	x	y	z	u
1	2.5909774	−0.937594	1.2060150	3.8180451
2	−56.179657	48.64028	4.2019976	140.7126609
3	−18.1669171	23.8983283	1.235399	70.8760384
4	−31.0773581	11.5813421	−0.0858966	35.8334771
5	−11.2341993	5.5159472	−0.5424367	18.1399774
6	−7.2376635	2.6570918	−0.5949845	8.9455653
7	−5.068011	0.8643791	0.0698393	7.463245
8	−3.1995777	0.8058214	−0.016966	2.5706599
9	−2.3006812	0.947281	−0.1259898	1.5493405
10	−1.7984976	1.2948349	−0.3986781	0.6413719
11	−1.7228517	1.3961188	−0.7257321	0.5096946
12	−1.6899483	1.4075929	−0.7773826	0.4861448
13	−1.6882143	1.40864	−0.7804767	0.4844253
14	−1.6882091	1.4086429	−0.7804865	0.4844206
15	−1.6882091	1.4086429	−0.7804865	0.4844206

Así, una de las raíces del sistema corresponde a $x = -1.6882091$, $y = 1.4086429$, $z = -0.7804865$, $u = 0.4844206$. Partiendo de los valores $x = y = z = u = -1$, se obtienen los siguientes resultados que conducen a otra raíz del sistema.


1
Precisión deseada de las raíces
N
i = 1
1
N
F[i]
j = 1
JF[i, j]
t = 0
c = 1
P[t, c]
t = t + 1
2
c = 1
Evaluar el vector B = F[i] en el punto P_i que es la matriz P[t − 1, c]
j = 1
Evaluar la matriz Jacobiana A = J[i, j] en P[t − 1, c]
JF[i, j] es singular
sí
no
Procedimiento de Gauss
Se resuelve el sistema lineal de N ecuaciones con n incógnitas AX = B, en donde A es la matriz JF[i, j] evaluada en P[t − 1, c] y B es el vector F[i] evaluado en P[t − 1, c], para c = 1, 2, . . . , N. La solución X es el vector V[j]
P[t, j] = P[t − 1, j] − V[j]
P[t, j]
Continuar
llegaste
t
F
Los productos de las iteraciones son P[1, j]
Captura de datos
Evaluación de la matriz n × 1 B y de la matriz Jacobiana A = J[i, j]
Procedimiento de Gauss
Imprime resultados

iteración	x	y	z	u
1	0.6666667	−1.6666667	−1.8333333	−1.1666667
2	0.286648	−1.48753	−1.1142345	−1.353069
3	0.2943675	−1.4977126	−1.0030817	−1.5343253
4	0.3010535	−1.4979912	−1.0074584	−1.5375728
5	0.301041	−1.4979962	−1.0074559	−1.5375777
6	0.301041	−1.4979962	−1.0074559	−1.5375777

Entonces otra raíz del sistema es $x = 0.301041$, $y = -1.4979962$, $z = -1.0074559$, $u = -1.5375777$. ■

Ejercicios (Capítulo 3, Sección 7)

1. Considere el sistema

$$(x^2 + y^2)e^{-(x^2+y^2)} - z = 0$$
$$Ay - z = 0$$
$$x^2 + y^2 + z^2 - 1 = 0$$

Resuelva el sistema para $A = 3$ y para $A = 1$. Verifique en cada caso que se obtienen dos soluciones del tipo $(\pm x_0, y_0, z_0)$. Para $A = A_0 = 0.391226858$, se obtiene una solución del tipo $(0, y_0, z_0)$ (con exactitud hasta cienmilésimas). Interprete geométricamente este problema. ¿Qué pasa para valores de A menores que A_0?

2. Sea $f: \mathbb{R}^3 \to \mathbb{R}$ la función

$$f(x, y, z) = x^2 + y^2 + z^2 + xy + xz + yz + 6x + 8y + 7z + 3$$

Resuelva (para x, y, z) la ecuación $\operatorname{grad} f(x, y, z) = f(x, y, z)(1, 1, 1)$.

3. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = (1 + y^2)\ln\cosh x + 2x^2 + 3y^2 - x - y + 1$. Resuelva la ecuación, para x, y, $\operatorname{grad} f(x, y) = \operatorname{grad}\left(\frac{1}{f}\right)(x, y)$.

En los ejercicios 4–10 resuelva el sistema dado.

4. $x^3 + y^3 - 2xy = 0, \quad x^2 + y^2 = 1$

5. $2x^4 - 2x^2y + y^2 - 2y^3 + y^4 = 0, \quad 3x^2 - 2xy - 6y^2 + 3 = 0$

6. $x^2 + y^2 - z - 2 = 0, \quad x^2 - y^2 - z = 0, \quad 3x + 2y - 6z = 0$

7. $5\operatorname{sen} xy + z - 1 = 0, \quad x + y - z = 0, \quad 3x^2 + 4y^2 + 6z^2 + xy + xz - yz + x - y + z - 5 = 0$

8. $\sqrt{\cos x^2 + y^2} - 6z^2 + 1 = 0 \quad xyz - 4 = 0, \quad e^x - e^y + e^z - 3 = 0$

9. $x^3 + y^3 + z^3 = 0, \quad \operatorname{sen} x + \operatorname{sen} y + \operatorname{sen} z + 1 = 0, \quad 4xy + 3x^2 + 5z^2 = 0$

10. $\operatorname{sen}(xy) + z - 1 = 0, \quad 2\operatorname{sen}(xz) + y - 2 = 0, \quad 3\operatorname{sen}(yz) + x - 3 = 0$

Capítulo



Extremos de las funciones de varias variables

En este capítulo estudiaremos lo referente a los máximos y mínimos de las funciones de varias variables. Antes de empezar recordemos rápidamente el tratamiento de este tema en el caso de funciones de una sola variable, así como los resultados más importantes del primer curso de cálculo. Siendo éste uno de los más importantes temas del curso —por sus aplicaciones, entre otros aspectos— esta pequeña retrospectiva al caso "simple" de funciones de una variable, nos servirá de guía para abordar de la misma manera la problemática presentada con funciones de varias variables (la cual, como es natural esperar, tiene algunas complicaciones técnicas adicionales).

La función $f: I \subseteq \mathbb{R} \to \mathbb{R}$, definida en el intervalo abierto I de $\mathbb{R}$, se decía tener un máximo (mínimo) local o relativo en un punto $x_0 \in I$ si en una vecindad V_{x_0} de x_0 se tenía $f(x_0) \geq f(x)$ ($f(x_0) \leq f(x)$, respectivamente) para toda x en V_{x_0}. En otras palabras, f tiene un máximo (mínimo) local en x_0 si $f(x_0)$ es el valor más grande (más pequeño) de la función en torno a x_0.

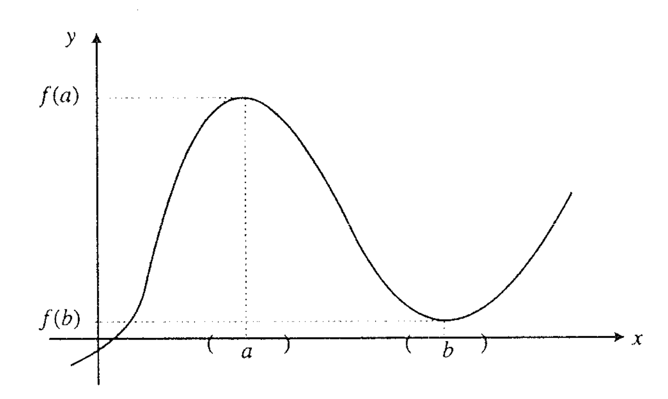


Figura 1. $y = f(x)$ tiene un máximo local en $x = a$ y un mínimo local en $x = b$.

Una condición necesaria para que la función f tenga un extremo (máximo o mínimo) local en x_0 es que, si $f'(x_0)$ existe, entonces $f'(x_0) = 0$. Por ejemplo, la función $f(x) = |x|$ tiene un mínimo local en $x = 0$, pues $f(0) = |0| = 0 \leq |x| = f(x)$ para toda x en una vecindad de $x = 0$. En este caso $f'(0)$ no existe. También la función $f(x) = -x^2$ tiene un máximo local en $x = 0$,

pues $f(0) = 0 \geq -x^2 = f(x)$ para toda x en una vecindad de $x = 0$. En este caso la función es diferenciable en $x = 0$ y se tiene $f'(0) = -2(0) = 0$. Geométricamente esta condición necesaria para la existencia de un extremo local, nos dice que si la gráfica de la función es suave en x_0, su recta tangente en ese punto debe ser horizontal (tal recta tangente podría no existir y la función tener extremo en x_0, como es el caso de $f(x) = |x|$ en $x = 0$).

A un punto $x_0 \in I$ en el que $f'(x_0)$ es cero o no existe, se le llama *punto crítico* de la función f. Entonces, una condición necesaria para que la función f tenga un extremo local en x_0 es que x_0 sea un punto crítico de ella. Tal condición no es suficiente, ya que, por ejemplo, la función $f(x) = x^3$ es tal que $f'(0) = 3(0)^2 = 0$ pero no tiene extremo local en $x = 0$ (pues en cualquier vecindad de $x = 0$, digamos $V_0 = \{x \in \mathbb{R}, -\epsilon < x < \epsilon\}$ se tiene para $0 \leq x < \epsilon$ que $f(x) = x^3 \geq 0 = f(0)$, en tanto que para $-\epsilon < x \leq 0$ se tiene $f(x) = x^3 \leq 0 = f(0)$).

Las condiciones suficientes para la existencia de extremo local de una función f en un punto x_0, son condiciones que garantizan la permanencia del signo de la expresión $f(x) - f(x_0)$ en los alrededores de x_0. Así, si tales condiciones nos aseguran que $f(x) - f(x_0)$ es no negativo en los alrededores de x_0 podríamos concluir que la función f tiene un mínimo local en x_0, en tanto que si $f(x) - f(x_0)$ es no positivo en los alrededores de x_0, se puede concluir la existencia de un máximo local para f en x_0.

Con la ayuda de la fórmula de Taylor para la función f en x_0 se pueden establecer tales condiciones suficientes. Siendo f suficientemente diferenciable en x_0, la fórmula de Taylor nos dice que podemos escribir, para h pequeñas (tal que $x_0 + h$ quede en el dominio de f)

$$f(x_0 + h) - f(x_0) = f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + r(h)$$

donde el residuo $r(h)$ tiene la propiedad de que $\lim_{h\to 0} \frac{r(h)}{h^2} = 0$. Es decir, $r(h)$ es más pequeño que h^2. Si x_0 es un punto crítico de f la expresión anterior se ve como

$$f(x_0 + h) - f(x_0) = \frac{1}{2}f''(x_0)h^2 + r(h)$$

Cuando h es pequeño, $x_0 + h$ es un punto en los alrededores de x_0 y, como $r(h)$ es muy pequeño en relación a $\frac{1}{2}f''(x_0)h^2$, el signo del segundo miembro de esta última expresión queda determinado por el del término $\frac{1}{2}f''(x_0)h^2$. Más aún, como $h^2 \geq 0$, el signo de $f(x_0 + h) - f(x_0)$ queda determinado por el signo de $f''(x_0)$, teniéndose que:

a. si $f''(x_0) > 0$, entonces $f(x_0 + h) - f(x_0) \geq 0$ para h pequeño (i.e. $f(x) - f(x_0) \geq 0$ para x en los alrededores de x_0) y por lo tanto f tendrá un mínimo local en x_0;

b. si $f''(x_0) < 0$, entonces $f(x_0 + h) - f(x_0) \leq 0$ para h pequeño, y por lo tanto f tendrá un máximo local en x_0.

c. si $f''(x_0) = 0$, la fórmula de Taylor anterior no da información suficiente para concluir la existencia de extremo local de f en x_0. Este es el llamado "criterio de la segunda derivada" para analizar el carácter de puntos críticos de funciones, el cual da condiciones suficientes para la existencia de extremos locales.

La forma de estudiar de los extremos locales de funciones de varias variables será similar a la breve exposición que acabamos de presentar para funciones de una variable. Por medio de la fórmula de Taylor se establecerán las condiciones suficientes para que una función tenga extremo local en un punto. En la primera sección se dan las definiciones básicas y se presentan los primeros

ejemplos que nos situarán en los problemas específicos que se tienen con los extremos locales de funciones de varias variables. En la segunda sección se estudia la fórmula de Taylor que nos servirá para, en la tercera sección, establecer el teorema que da condiciones suficientes para la existencia de extremos locales. En la cuarta sección presentamos un teorema para el caso concreto de funciones de dos variables, y dedicamos el resto de la sección a resolver algunos ejemplos típicos de aplicación de la teoría previamente estudiada. Por último en las quinta y sexta secciones se estudia la importante problemática de los extremos sujetos a restricciones, en las que se discute el método de los multiplicadores de Lagrange para atacar este tipo de problemas.

4.1 Definiciones y ejemplos preliminares

Definición. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$. Se dice que f tiene un máximo (mínimo) local o relativo en el punto $\mathbf{x}_0 \in U$ si $f(\mathbf{x}_0) \geq f(\mathbf{x})$ ($f(\mathbf{x}_0) \leq f(\mathbf{x})$ respectivamente) para toda $\mathbf{x}$ en una bola B de centro en $\mathbf{x}_0$. ∎

Es decir, al igual que en el caso de funciones de una variable, la función f tendrá un máximo (mínimo) local en $\mathbf{x}_0 \in U$ si $f(\mathbf{x}_0)$ es el valor más grande (más pequeño, respectivamente) de todos los valores de $f(\mathbf{x})$ para $\mathbf{x}$ en una bola de centro en $\mathbf{x}_0$

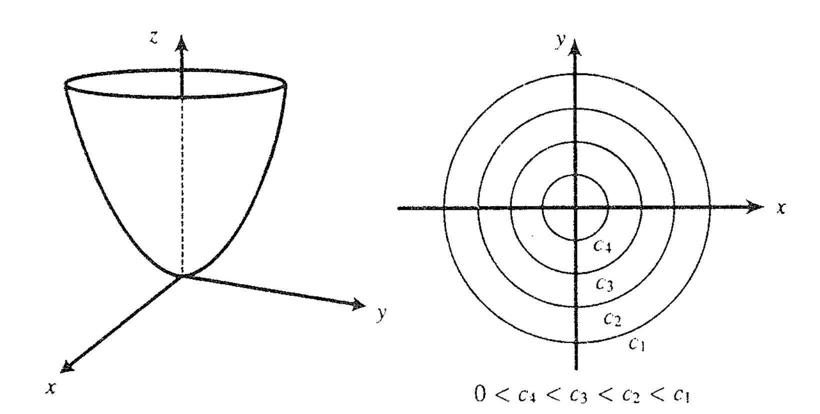


Figura 1. Gráfica de la función $z = x^2 + y^2$ y sus curvas de nivel cerca del origen.

Ejemplo 1. La función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = x^2 + y^2$$

tiene un mínimo local en $\mathbf{x}_0 = (0, 0)$ pues en una bola B de centro en $\mathbf{x}_0$ (de hecho, en cualquier bola B de centro en $\mathbf{x}_0$) se tiene

$$f(0, 0) = 0^2 + 0^2 = 0 \leq x^2 + y^2 = f(x, y)$$

para toda $(x, y) \in B$. Recuerde el aspecto geométrico de la superficie $z = x^2 + y^2$. Su gráfica cerca del origen muestra el comportamiento típico que una superficie $z = f(x, y)$ presenta en un punto de

mínimo local. Más aún, observe que las curvas de nivel de esta función son $x^2 + y^2 = c$ $(c > 0)$, cuyo aspecto, a medida que c es más cercano a cero, es el de un conjunto de círculos concéntricos cada vez "más pegados" (este es el aspecto típico del conjunto de curvas de nivel de una función con un extremo local). ■

Ejemplo 2. La función $f: \mathbb{R}^2 \to \mathbb{R}$, dada por

$$f(x, y) = 1 - x^2 - y^2$$

tiene un máximo local en $\mathbf{x}_0 = (0, 0)$, pues en una bola B de centro en $\mathbf{x}_0$ se tiene

$$f(0, 0) = 1 \geq 1 - x^2 - y^2 = f(x, y)$$

para toda $(x, y) \in B$. Nuevamente llamamos la atención al aspecto geométrico de la superficie $z = f(x, y)$, así como el de sus curvas de nivel cerca del origen.

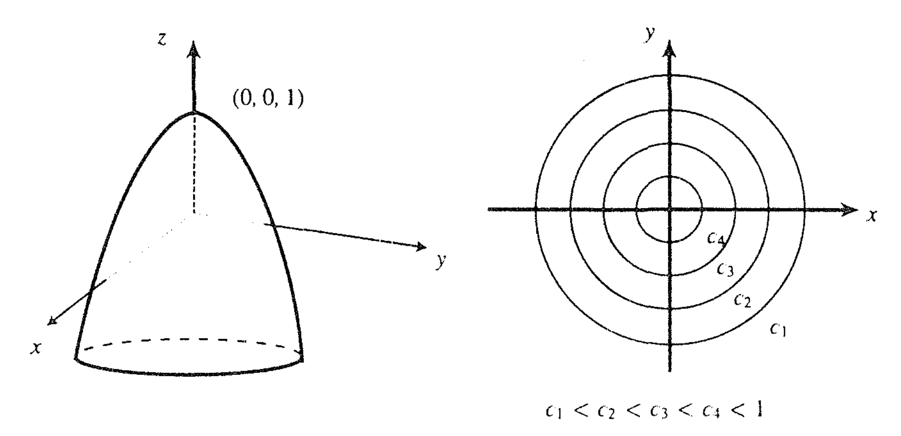


Figura 2. Gráfica de la función $z = 1 - x^2 - y^2$ y sus curvas de nivel. ■

Ejemplo 3. Los dos ejemplos anteriores muestran funciones diferenciables que tienen extremos locales. Una función puede no ser diferenciable en un punto y tener ahí un extremo local (lo mismo que en el caso de funciones de una variable: $f(x) = |x|$ en $x = 0$). Por ejemplo la función $f: \mathbb{R}^2 \to \mathbb{R}$, dada por $f(x, y) = \sqrt{x^2 + y^2}$ tiene un mínimo local en (0,0) pues

$$f(0, 0) = 0 \leq \sqrt{x^2 + y^2} = f(x, y)$$

para toda (x, y) en una bola (cualquiera) B de centro en $(0, 0)$. Sin embargo tal función no es diferenciable en $(0, 0)$ pues sus derivadas parciales no se presentan ahí. La superficie $z = \sqrt{x^2 + y^2}$ es un cono abierto hacia arriba con vértice en el origen. El conjunto de curvas de nivel cerca del origen tiene el mismo aspecto que el de la superficie $z = x^2 + y^2$ del ejemplo 1. ■

Supongamos que la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $\bar{\mathbf{x}} \in U$ y que ahí tiene un extremo local. Digamos que $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$. Para cada $i = 1, 2, \ldots, n$ considere la función

$\varphi_i: I \subseteq \mathbb{R} \to \mathbb{R}$, definida en alguna vecindad de $\bar{x}_i$, digamos $I = \{x \in \mathbb{R} | \bar{x}_i - \epsilon < x < \bar{x}_i + \epsilon\}$ (ϵ puede ser el radio de la bola B de la que habla la definición de extremo local de f), dada por

$$\varphi_i(x) = f(\bar{x}_1, \ldots, \bar{x}_{i-1}, x, \bar{x}_{i+1}, \ldots, \bar{x}_n)$$

Es claro que teniendo f un extremo local en $\bar{\mathbf{x}}$, la función φ_i debe tener también un extremo local (del mismo tipo) en $\bar{x}_i$, de modo que $\varphi_i'(\bar{x}_i) = 0$. Es decir que

$$\frac{\partial f}{\partial x_i}(\bar{\mathbf{x}}) = 0, \quad i = 1, 2, \ldots, n$$

Así, una condición *necesaria* para que la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en $\bar{\mathbf{x}} \in U$, tenga en ese punto un extremo local es que todas sus derivadas parciales se anulen en $\bar{\mathbf{x}}$.

Definición. Al punto $\bar{\mathbf{x}} \in U$ en el que todas las derivadas parciales de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ se anulan, se le llama *punto crítico* (o punto estacionario) de la función. ■

Considere por ejemplo una función de dos variables $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$. Suponga que en el punto (x_0, y_0) esta función tiene un extremo local. Si f es diferenciable en (x_0, y_0) se debe tener que

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{y} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

Obsérvese entonces que la ecuación del plano tangente a $z = f(x, y)$ en (x_0, y_0) es $z = f(x_0, y_0)$, el cual es un plano horizontal (paralelo al plano xy). Entonces, desde el punto de vista geométrico, una condición necesaria para que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ diferenciable en $\mathbf{p} \in U$, tenga ahí un extremo local, es que en ese punto su plano tangente sea horizontal. Sin embargo, como era de esperarse, tal condición está muy lejos de ser suficiente.

Ejemplo 4. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = x^2 - y^2$. Esta función es diferenciable en todo $\mathbb{R}^2$. Los puntos críticos de f se obtienen al resolver el sistema

$$\frac{\partial f}{\partial x} = 2x = 0, \qquad \frac{\partial f}{\partial y} = -2y = 0$$

de donde se ve que $(0, 0)$ es el único punto crítico que existe. Así, si la función f tiene algún extremo local, lo tiene en $(0, 0)$. Sea B una bola con centro en el origen, digamos $B = \{(x, y) | x^2 - y^2 < \epsilon\}$. Los puntos del tipo $(r, 0)$ con $r^2 < \epsilon^2$, están en B y para ellos se tienen

$$f(0, 0) = 0 \leq r^2 = f(r, 0)$$

Del mismo modo, los puntos del tipo $(0, r)$ con $r^2 < \epsilon^2$, está en B y para ellos se tiene

$$f(0, 0) = 0 \geq -r^2 = f(0, r)$$

de modo que la expresión: $f(x, y) - f(0, 0)$ no mantiene signo constante en ninguna bola con centro en el origen. Es decir, la función f no tiene extremo local en $(0, 0)$. ■

Los puntos que tienen un comportamiento como el origen en el ejemplo anterior reciben un nombre especial.

Definición. Considere la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$. Sea $\bar{\mathbf{x}} \in U$. Si cualquier bola B (de $\mathbb{R}^n$) con centro en $\bar{\mathbf{x}}$ contiene puntos $\mathbf{x} \in B$ tales que $f(\mathbf{x}) - f(\bar{\mathbf{x}})$ es positivo y puntos $\mathbf{y} \in B$ tales que $f(\mathbf{y}) - f(\bar{\mathbf{x}})$ es negativo, se dice que $\bar{\mathbf{x}}$ es un *punto de ensilladura* (o *punto silla*) de la función f. ∎

El aspecto de la gráfica de una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ cerca de un punto de ensilladura, es justamente como si ésta fuera una silla de montar. Así lo muestra la función $f(x, y) = x^2 - y^2$ del ejemplo 4, cuya gráfica cerca del origen se muestra a continuación así como algunas de sus curvas de nivel

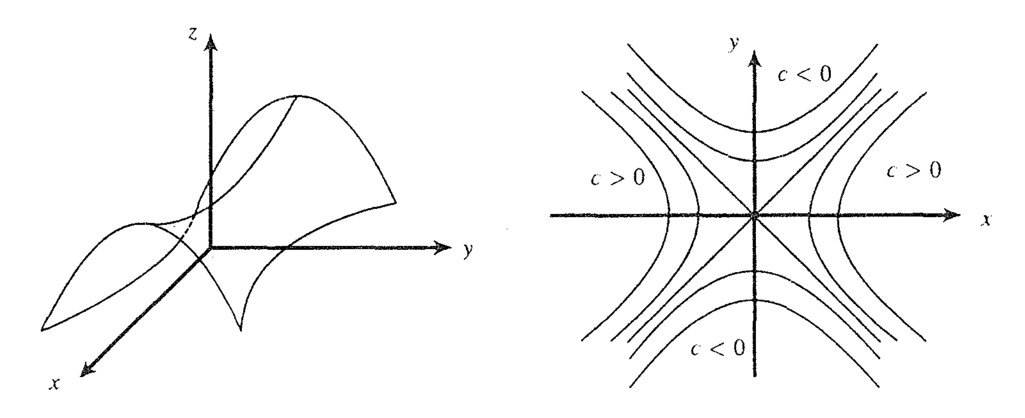


Figura 3. Gráfica y curvas de nivel de la función $z = x^2 - y^2$.

Nótese que las curvas de nivel delatan el por qué esta función no puede tener extremo local en el origen. Cualquier bola con centro en el origen contendrá puntos de curvas con $c > 0$ (donde la superficie está por encima del plano xy) y puntos de curvas con $c < 0$ (donde la superficie está por debajo del plano xy).

Un acercamiento natural hacia la obtención de condiciones suficientes para la existencia de extremo local de una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto crítico $\bar{\mathbf{x}}$ de ella, sería estudiar las funciones $\varphi_i(x) = f(\bar{x}_1, \ldots, \bar{x}_{i-1}, x, \bar{x}_{i+1}, \ldots, \bar{x}_n)$ mencionadas anteriormente, a las cuales podemos aplicar (con f suficientemente diferenciable) el criterio de la segunda derivada para funciones de una variable. Parecería lógico pensar que si la función $\varphi_i(x)$ tiene un máximo local (o mínimo local) en $x = \bar{x}_i$, para toda $i = 1, 2, \ldots, n$, entonces la función f deberá tener un máximo local (o mínimo local) en el punto $\bar{\mathbf{x}}$. Para una función de dos variables esta situación se ve así: si $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ tiene un punto crítico en $(\bar{x}, \bar{y})$, consideramos las funciones $\varphi_1(x) = f(x, \bar{y})$, $\varphi_2(y) = f(\bar{x}, y)$. Obsérvese que $\varphi_1(x)$ geométricamente representa la curva de intersección de la superficie $z = f(x, y)$ con el plano $y = \bar{y}$, y $\varphi_2(y)$ la curva de intersección de $z = f(x, y)$ con el plano $x = \bar{x}$. Supongamos que estas dos curvas tienen máximos (o mínimos) locales (en $x = \bar{x}$ y $y = \bar{y}$, respectivamente). ¿Se infiere de aquí que la función $z = f(x, y)$ tiene un máximo (o mínimo) local en $(\bar{x}, \bar{y})$? La respuesta a esta pregunta es *NO*. Ciertamente tal condición es necesaria, pero no suficiente.

Antes de ver un ejemplo concreto que muestre lo anterior, vea que la función $z = f(x, y) = x^2 - y^2$ no cumple con tal condición necesaria. En efecto, el punto $(0, 0)$ es el único punto crítico. La función $\varphi_1(x) = f(x, 0) = x^2$ tiene un mínimo local en $x = 0$ en tanto que la función $\varphi_2(y) = f(0, y) = -y^2$ tiene un máximo local en $y = 0$.

Ejemplo 5. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = 1 - x^2 - y^2 + 3xy$. Esta es una función diferenciable en todo $\mathbb{R}^2$. Sus puntos críticos se obtienen al resolver el sistema

$$\frac{\partial f}{\partial x} = -2x + 3y = 0, \qquad \frac{\partial f}{\partial y} = -2y + 3x = 0$$

Se ve que $x = 0$, $y = 0$ es la única solución, por lo que $(0, 0)$ es el único punto crítico de f. Considere las funciones $\varphi_1(x) = f(x, 0) = 1 - x^2$, $\varphi_2(y) = f(0, y) = 1 - y^2$. Ambas tienen un máximo local en $x = 0$ y $y = 0$ (respectivamente). Sin embargo, la función f no tiene un máximo local en $(0, 0)$. En efecto, basta que veamos que en cualquier bola B con centro en el origen, hay puntos $(x, y) \in B$ para los que $f(x, y) \geq f(0, 0) = 1$. Sea B una bola cualquiera con centro en $(0, 0)$, digamos que $B = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < \epsilon^2\}$. Considere los puntos $(\xi, \xi) \in B$ (con $\xi^2 < \frac{1}{2}\epsilon^2$). Observe que $f(\xi, \xi) = 1 - \xi^2 - \xi^2 + 3(\xi)(\xi) = 1 + \xi^2 \geq 1 = f(0, 0)$. Así pues, en la bola B hay puntos (x, y) para los cuales $f(x, y) - f(0, 0)$ es negativo (los puntos del tipo $(x, 0)$ ó $(0, y)$ $x \neq 0$ y $y \neq 0$) y también puntos en los cuales $f(x, y) - f(0, 0)$ es positivo (los del tipo (x, x) $x \neq 0$). Se trata pues de un punto de ensilladura. Desde el punto de vista geométrico, acontece entonces que la superficie $z = 1 - x^2 - y^2 + 3xy$ es tal que su intersección con los planos $x = 0$ y $y = 0$ produce curvas con máximo local en el origen, pero su intersección con el plano $y = x$, produce una curva ($z = 1 + x^2$) con un mínimo local en el origen. ■

El ejemplo anterior nos invita a considerar, en la misma dirección de "tipo de extremos de las curvas de corte de la superficie $z = f(x, y)$ con planos perpendiculares al plano xy", condiciones más fuertes a partir de las cuales podamos concluir la existencia de un extremo local para la función dada. Sería natural ahora preguntarse lo siguiente: si la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ la cual es diferenciable en el punto crítico $(x_0, y_0) \in U$ es tal que todas las curvas procedentes de la intersección la superficie $z = f(x, y)$ con planos perpendiculares al plano xy que pasen por (x_0, y_0), tiene el mismo tipo de extremos en (x_0, y_0), ¿tiene la función f misma un extremo local de la misma naturaleza en ese punto? Más aún, supongamos que al intersectar la superficie $z = f(x, y)$, que tiene un punto crítico en (x_0, y_0), con el plano $y = k(x - x_0) + y_0$, se obtiene una curva $z = f(x, k(x - x_0) + y_0)$ la cual tiene un máximo local, digamos, en $x = x_0$. ¿Se concluye de aquí que la función misma $z = f(x, y)$ tiene un máximo local en (x_0, y_0)? Nuevamente observamos que tal condición es ciertamenete necesaria para la existencia del extremo local (la cual no se satisface por la función del ejemplo 5). Y aunque ahora sí parecería que la condición es lo suficientemente fuerte como para garantizar la existencia del extremo local para la función f, el siguiente ejemplo nos muestra que con funciones de varias variables pueden ocurrir situaciones extrañas en las que la intuición geométrica nos falla.

Ejemplo 6. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = 2x^2 - 3xy^2 + y^4$, la cual es diferenciable en todo $\mathbb{R}^2$. Las derivadas parciales de esta función son

$$\frac{\partial f}{\partial x} = 4x - 3y^2, \qquad \frac{\partial f}{\partial y} = -6xy + 4y^3$$

de donde se ve que $(0, 0)$ es un punto crítico de f. Tomemos los planos $y = kx$, los cuales pasan por el origen y son perpendiculares al plano xy. La intersección de la superficie $z = 2x^2 - 3xy^2 + y^4$ con el plano $y = kx$ es $z = \varphi(x) = 2x^2 - 3x(kx)^2 + (kx)^4 = 2x^2 - 3k^2x^3 + k^4x^4$. Se tiene $\varphi'(x) = 4x - 9k^2x^2 + 4k^4x^3$, de donde $\varphi'(0) = 0$, como tenía que ocurrir. Más aún, como $\varphi''(x) = 4 - 18k^2x + 12k^4x^2$, se tiene $\varphi''(0) = 4 > 0$, por lo que la curva $\varphi(x)$ tiene en el origen un *mínimo* local. Tenemos entonces que nuestra superficie $z = f(x, y)$ es tal que

"rebanándola" con planos perpendiculares al plano xy que pasan por el origen, las correspondientes curvas de intersección tienen un mínimo en $x = 0$. Sin embargo, veamos que la función f *no tiene* un mínimo en el origen. Si así fuera, se debería de cumplir que $f(x, y) \geq 0 = f(0, 0)$ para toda (x, y) en una bola con centro en $(0, 0)$, digamos $B = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < \epsilon^2\}$. Veamos que existen puntos en esta bola para los que $f(x, y) \leq 0$. Si escribimos a $f(x, y)$ como $f(x, y) = (x - y^2)(2x - y^2)$, vemos que $(x - y^2)(2x - y^2) \leq 0$ tiene solución para $x \leq y^2 \leq 2x$. Tomando puntos suficientemente cerca del origen que satisfagan esta desigualdad, por ejemplo puntos de la forma $(\frac{2}{3}\xi^2, \xi)$, en donde $\frac{4}{9}\xi^4 + \xi^2 \leq \epsilon^2$ (esto garantiza que todos los puntos están en la bola B), se tiene $f(\frac{2}{3}\xi^2, \xi) = (\frac{2}{3}\xi^2 - \xi^2)(\frac{4}{3}\xi^2 - \xi^2) = -\frac{1}{4}\xi^2 \leq 0$ por lo que la función f tiene, en realidad, un punto de ensilladura en el origen.

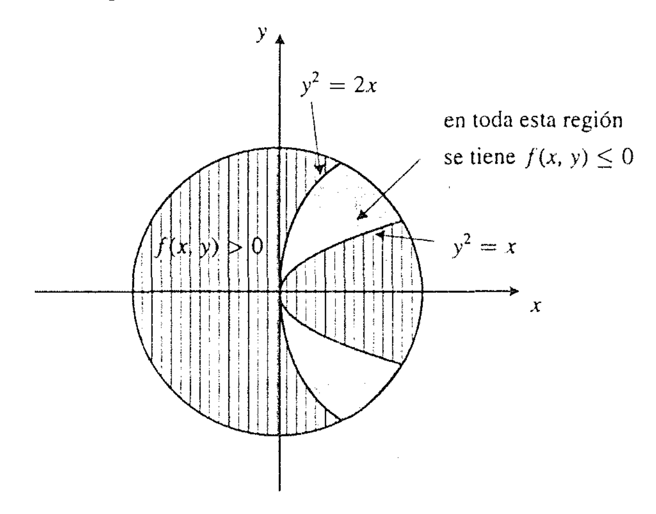


Figura 4. Diagrama de la bola B en la que $f(x, y) < 0$. ■

Acontece que las condiciones suficientes que garantizan la existencia de un extremo local de una función f en un punto crítico de la misma se tendrán que plantear, al igual que el caso de una variable, con la ayuda de la fórmula de Taylor para la función f en un punto crítico. En la siguiente sección estudiaremos lo referente al Teorema de Taylor, que establece la fórmula mencionada.

Ejercicios (Capítulo 4, Sección 1)

1. Directamente de la definición, pruebe que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = ax^2 + by^2$, donde a y b son reales positivos dados, tiene un mínimo local en el origen. (Ver ejemplo 1).

2. Generalizando el ejercicio anterior, pruebe que la función $f: \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = a_1(x_1)^i 1 + a_2(x_2)^i 2 + \ldots + a_n(x_n)^i n$, donde los coeficientes a_i son reales positivos dados y los exponentes i_j son enteros positivos pares, tiene un mínimo local en el origen.

3. Pruebe que si la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, definida en el conjunto abierto U de $\mathbb{R}^n$, tiene un máximo local en el punto $\mathbf{p} \in U$, entonces la función $-f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $(-f)(\mathbf{x}) = -f(\mathbf{x})$, tiene un mínimo local en ese punto.

4. Demuestre que si la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, definida en el conjunto abierto U de $\mathbb{R}^n$, tiene un máximo (mínimo) local en el punto $\mathbf{p} \in U$, entonces la función $F: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $F(\mathbf{x}) = f(\mathbf{x}) + k$, donde k es un número real dado, tiene también un máximo (mínimo) local en $\mathbf{p}$. Interprete geométricamente este hecho. (Ver ejercicio 1 de la sección 2, capítulo 2).

5. Sean $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones definidas en el conjunto abierto U de $\mathbb{R}^n$. Suponga que estas funciones tienen un máximo (mínimo) local en el punto $\mathbf{p} \in U$. Demuestre que la función suma $F: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ tiene un máximo (mínimo) local en $\mathbf{p}$. ¿Podemos concluir lo mismo con la función producto $F(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$? Explique.

6. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = ax^2 + by^2$, en que a y b son números reales tales que $ab < 0$, tiene un punto de ensilladura en el origen, mostrando que en cualquier bola $B = \{(x, y) | x^2 + y^2 < \epsilon^2\}$ existen puntos $\mathbf{p}, \mathbf{q} \in B$, tales que $f(\mathbf{p}) > 0$ y $f(\mathbf{q}) < 0$. (Sugerencia: procure puntos en los ejes coordenados).

7. Generalizando el ejercicio anterior: demuestre que la función $f: \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$, donde no todos los coeficientes a_i tienen el mismo signo, tiene un punto de ensilladura en el origen.

8. Demuestre que la función $f(x, y) = x^3 + y^3$ tiene un punto de ensilladura en el origen.

9. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = |x| + |y|$, tiene un mínimo local en el origen. Observe que f no es diferenciable en $(0, 0)$ ¿por qué? Haga un esquema mostrando las curvas de nivel de esta función cerca del origen.

10. Generalizando el ejercicio anterior: demuestre que la función $f: \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = a_1|x_1| + a_2|x_2| + \ldots + a_n|x_n|$, donde todos los coeficientes a_i son positivos, tiene un mínimo local en el origen.

11. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^3 + 3x + y^3 + y$ no tiene extremos locales. (Sugerencia: si los tuviera, éstos deberían ser puntos críticos de f ...).

12. Demuestre que la función $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^5 + y^5 + z^7 + 4x + 2y + 9z + 3$ no tiene extremos locales.

13. Demuestre que la función $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = \arctan(x + 2y + 3z)$ no tiene extremos locales.

14. Demuestre que una función lineal $f: \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$ no puede tener extremos locales.

15. Una función constante $f: \mathbb{R}^n \to \mathbb{R}$, $f(\mathbf{x}) = c$, tiene extremos locales?

16. Demuestre que si la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + 3xy + y^2 + 2x - 4y + 210$ tiene un extremo local, éste debe de estar en el punto $\mathbf{p} = (16/5, -14/5)$.

17. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = ax^2 + 2bxy + cy^2 + dx + ey + f$. Demuestre que esta función puede tener a lo más un extremo local, y que si lo tiene, entonces $ac - b^2 \neq 0$. Use el ejercicio 6 para probar que la condición $ac - b^2 \neq 0$ no garantiza la existencia de extremo local de f.

18. Considere la función $f: \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = \ln(1 + x_1^2 + x_2^2 + \ldots + x_n^2)$. Demuestre que esta función puede tener a lo más un extremo local, el cual se encontraría en el origen. Básese en que la función logaritmo es creciente en todo su dominio para probar que *de hecho* existe tal extremo local en el origen. ¿Qué tipo de extremo es?

19. Sea $f: \mathbb{R}^+ \to \mathbb{R}$ una función diferenciable con derivada siempre positiva. Demuestre que la función $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = f(1 + x^2 + y^2)$ tiene un mínimo local en el origen. Más en general, demuestre que la función $F: \mathbb{R}^n \to \mathbb{R}$, $F(x_1, x_2, \ldots, x_n) = f(1 + x_1^2 + x_2^2 + \ldots + x_n^2)$ tiene un mínimo local en el origen. Nótese que el ejercicio anterior es un caso particular de este ejercicio.

20. Suponga que la función f del ejercicio anterior tiene derivada siempre negativa. Demuestre que la función F tiene un máximo local en el origen. Use este hecho para probar que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \frac{1}{1+x^2+y^2}$ tiene un máximo local en $(0, 0)$, y que lo mismo ocurre con la función $g(x, y) = e^{-1-x^2-y^2}$.

21. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^{2/3} + y^{2/3}$. Constate que el origen es un punto crítico de esta función. Use el hecho de que $f(x, y) \geq 0$ para todo $(x, y) \in \mathbb{R}^2$ para concluir que f tiene un mínimo local en $(0, 0)$.

22. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = (x^2 + y^2)^{1/3}$. Demuestre que f tiene un mínimo local en el origen.

23. Sea $f: \mathbb{R} \to \mathbb{R}$, $z = f(y)$, una función par que tiene un extremo local en el punto $(0, z_0)$. Considere la superficie de revolución $z = f\left(\sqrt{x^2 + y^2}\right)$. Demuestre que ésta tiene, en el origen, el mismo tipo de extremo local que la función f. Compruebe que el ejercicio anterior es un caso particular de la situación expuesta en este ejercicio.

24. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$ tiene un mínimo local en el origen. (Sugerencia: basta demostrar que la función $f(x) = x^2e^{-x^2}$ tiene un mínimo local para $x = 0$ y usar el resultado del ejercicio anterior).

25. Demuestre que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^4 + x^2y^2 + y^4$ tiene solamente un extremo local en el origen, el cual es un mínimo.

26. Generalizando el ejercicio anterior, demuestre que si $n = 4k$, $k \in \mathbb{N}$, entonces la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^n + cx^{n/2}y^{n/2} + y^n$, donde c es un número positivo dado, tiene solamente un extremo local, el cual es un mínimo.

27. Es fácil ver que si la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^ny^n$, donde n es un número natural dado, tiene extremos locales, éste es único y se debe encontrar en el origen (¿por qué?). Discuta la existencia y naturaleza de este extremo en términos de n.

En los ejercicios 28–40, encuentre los puntos críticos de la función dada. Recuerde que éstos son los puntos en que *puede* haber extremos locales de la función.

28. $f(x, y) = 3x + 8y - 2xy + 4$

29. $f(x, y) = x^2 + x + y^2 + 1$

30. $f(x, y) = x^2 + 2x + y^2 - 4y + 10$

31. $f(x, y) = 2x^3 + 3x^2 + 6x + y^3 + 3y + 12$

32. $f(x, y) = x^2y - x^2 - 3xy + 3x + 2y - 2$

33. $f(x, y) = x^2y^2 + x^2 - 5xy^2 - 5x + 6y^2 + 6$

34. $f(x, y) = (x - y)e^{x+2y}$

35. $f(x, y) = x \cos y$

36. $f(x, y) = x \cosh y$

37. $f(x, y, z) = 3x^4 - 8y^3 + 134z^{23} - 5$

38. $f(x, y, z) = xy + xz + yz - 3$

39. $f(x, y, z) = x + y + z + xy + xz + yz - 3$

40. $f(x, y, z) = xyz - 3xy - 2xz + 6x - yz + 3y + 2z - 6$

41. $f(x, y, z) = xyz^2 + xy + xz^2 + x - 2yz^2 - 2y - 2z^2 - 2$

42. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = 3x^4 - 4x^2y + y^2$. Compruebe que el punto $\mathbf{p} = (0, 0)$ es un punto crítico de esta función. Considere las curvas de intersección de la superficie $z = f(x, y)$ con planos que pasan por el origen, perpendiculares al plano xy. Demuestre que todas estas curvas tienen un mínimo local en el origen y que sin embargo la función f tiene un punto de ensilladura en el origen. (Sugerencia: use la mismas ideas del ejemplo 6).

43. Repita el ejercicio anterior con la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^6 - x^2y - x^4y + y^2$. Demuestre que, de hecho, tal situación se repite con la función $g: \mathbb{R}^2 \to \mathbb{R}$, $g(x, y) = x^{n+m} - x^m y - x^n y + y^2$, donde m y n son números pares.

44. Repita el ejercicio 42 con la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = y^2 + y - y\cos x + yx^2 + x^2 - x^2 \cos x$.

45. Los tres ejercicios anteriores son casos particulares de la siguiente situación más general. Sean f, $g: \mathbb{R} \to \mathbb{R}$ dos funciones diferenciables tales que en una bola con centro en el origen la gráfica de una de ellas siempre está por encima de la gráfica de la otra, y tales que tienen en el origen un máximo o un mínimo local (ambas el mismo tipo de extremo). Considere la función $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = y^2 - (f(x) + g(x))y + f(x)g(x)$. El origen es un punto crítico de F. La intersección de la superficie $z = F(x, y)$ con todos los planos perpendiculares al plano xy que pasan por el origen, son curvas que tienen en el origen el mismo tipo de extremo de f y g. Sin embargo, la función F tiene un punto de ensilladura en el origen. Demuestre estos hechos. Observe, sin embargo, que la situación mostrada en el ejemplo 6 es diferente de la aquí descrita. ¿Por qué?

46. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto U de $\mathbb{R}^n$, tal que $f(x) \neq 0\ \forall x \in \mathbb{R}^n$. Sea $\mathbf{p}$ un punto de U en el que

$$\operatorname{grad}\left(\frac{1}{f}\right)(\mathbf{p}) = \operatorname{grad} f(p)$$

Demuestre que $\mathbf{p}$ es un punto crítico de f. ¿Es cierta la afirmación recíproca?

4.2 La fórmula de Taylor de segundo orden

Si la función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ definida en el intervalo abierto I de $\mathbb{R}$ es derivable en $x_0 \in I$, entonces podemos escribir

$$f(x_0 + x) = f(x_0) + f'(x_0)x + r(x)$$

(para x tales que $x_0 + x \in I$) en que el residuo $r(x)$ tiene la propiedad de que $\lim_{x \to 0} \frac{r(x)}{x} = 0$. Una manera geométrica de leer esta fórmula (que no es más que la definición de derivada de f en x_0) es la siguiente: para valores pequeños de x se tiene la igualdad aproximada $f(x_0 + x) \approx f(x_0) + f'(x_0)x$ o bien $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$, la cual nos dice que si quisiéramos ver la gráfica de $y = f(x)$ en los alrededores del punto $(x_0, f(x_0))$ como una recta, ésta debería ser $y = f(x_0) + f'(x_0)(x - x_0)$, es decir, la recta que pasa por $(x_0, f(x_0))$ y tiene como pendiente a $f'(x_0)$ (que es justamente la recta tangente a $y = f(x)$ para $x = x_0$). Esta información resulta muy valiosa cuando investigamos algunos aspectos geométricos de la gráfica de la función $y = f(x)$. Por ejemplo, si al ver a $y = f(x)$ como su recta tangente en $(x_0, f(x_0))$, esta recta tiene pendiente positiva, entonces podemos decir que la función $y = f(x)$ es creciente en ese punto. Sin embargo, para otros aspectos esta información, que llamaremos "lineal" o "de primer orden", resulta insuficiente. Supongamos que estamos investigando los extremos locales de la función $y = f(x)$. Hemos detectado que en $x = x_0$ existe un punto crítico para la función (que suponemos es derivable). Se debe tener entonces $f'(x_0) = 0$. Si queremos ver como recta a esta función en los alrededores de $(x_0, f(x_0))$, esta recta sería (su recta tangente en ese punto) $y = f(x_0) - f'(x_0)(x - x_0) + f(x_0) = f(x_0)$. Esta información lineal lo único que nos dice es que en el punto $(x_0, f(x_0))$ la gráfica de $y = f(x)$ tiene una recta tangente horizontal (de ecuación $y = f(x_0)$), pero no nos dice nada acerca de la naturaleza del extremo en cuestión (si es que existe).

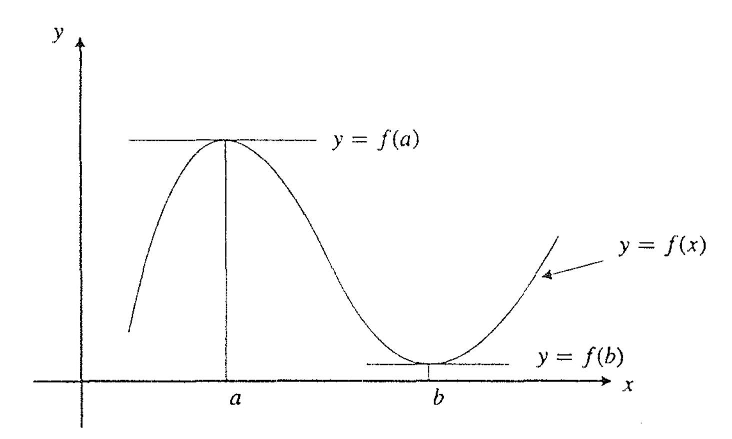


Figura 1. La recta tangente a $y = f(x)$ en un extremo local.

Se necesita entonces una información más fina para $y = f(x)$ en los alrededores de $(x_0, f(x_0))$ o sea: quisiéramos ver ya no como recta la gráfica de $y = f(x)$ en $(x_0, f(x_0))$, sino más bien, como una *parábola* con su vértice en ese punto. ("Querer ver como (recta-parábola) la gráfica de la función $y = f(x)$ en los alrededores de $(x_0, f(x_0))$" significa que la gráfica de la función $y = f(x)$ y la (recta-parábola) (las cuales obviamente deben pasar por el punto en cuestión) tengan, localmente, en los alrededores del punto $(x_0, f(x_0))$, las mismas propiedades geométricas; por ejemplo, la misma pendiente, concavidad, etc., lo cual se traduce en que en ese punto las derivadas de $y = f(x)$ y las de la (recta-parábola) coincidan). Si así fuera el caso, digamos, que tal parábola es $y = ax^2 + bx + c$. Como ésta debe pasar por $(x_0, f(x_0))$ la podemos escribir como $y = (x - x_0)^2 + b(x - x_0) + f(x_0)$.

Además queremos que su vértice esté en el punto $(x_0, f(x_0))$; entonces $y'(x_0) = 0$, de donde $b = 0$. Por último, queremos que la concavidad de la parábola sea la misma que la concavidad de la curva $y = f(x)$ para $x = x_0$ (ésta es la cuestión importante). Entonces $y''(x_0) = 2a = f''(x_0)$. Concluimos entonces que si la gráfica de la función $y = f(x)$ se viera como una parábola con vértice

en $(x_0, f(x_0))$, ésta será

$$y = \frac{1}{2} f''(x_0)(x - x_0)^2 + f(x_0) \tag{p}$$

(recuerde que la función tiene un punto crítico en x_0). De aquí se puede concluir lo siguiente (criterio de la segunda derivada para la existencia de extremos locales de la función $y = f(x)$ en el punto crítico x_0)

1. Si $f''(x_0) > 0$, la parábola (p) abre hacia arriba, y por lo tanto la función $y = f(x)$ tiene un mínimo local en $x = x_0$ (figura 2)

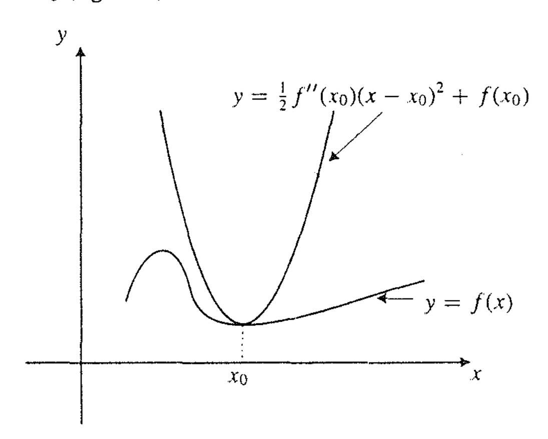


Figura 2. Un mínimo local en la función $y = f(x)$.

2. Si $f''(x_0) < 0$, la parábola (p) abre hacia abajo, y por lo tanto la función $y = f(x)$ tiene un máximo local en $x = x_0$ (figura 3)

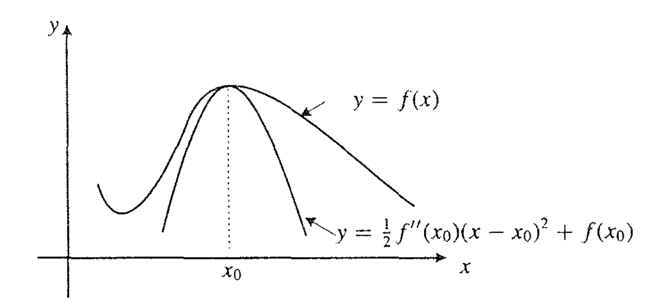


Figura 3. Un máximo local en la función $y = f(x)$

Vemos pues que esta información "cuadrática" o de "segundo orden" sí nos proporciona información acerca de la naturaleza de los puntos críticos en cuestión.

En general, si queremos ver la gráfica de $y = f(x)$ como una parábola alrededor del punto $(x_0, f(x_0))$ (donde ahora x_0 no es necesariamente un punto crítico de $y = f(x)$), ésta debería ser

$$y = \frac{1}{2} f''(x_0)(x - x_0)^2 + f'(x_0)(x - x_0) + f(x_0)$$

por lo que, para x cerca de x_0 podemos escribir la igualdad aproximada

$$f(x) \approx \frac{1}{2} f''(x_0)(x - x_0)^2 + f'(x_0)(x - x_0) + f(x_0)$$

o bien

$$f(x_0 + x) \approx f(x_0) + f'(x_0)x + \frac{1}{2} f''(x_0)x^2$$

Acontece que si f es una función con derivadas hasta segundo orden continuas vecinas a x_0, entonces podemos escribir

$$f(x_0 + x) = f(x_0) + f'(x_0)x + \frac{1}{2} f''(x_0)x^2 + r(x)$$

donde el residuo $r(x)$ tiene la propiedad de que $\lim_{x \to 0} \frac{r(x)}{x^2} = 0$. Una alternativa es

$$f(x_0 + x) = f(x_0) + f'(x_0)x + \frac{1}{2} f''(\xi)x^2 \qquad (*)$$

donde

$$x_0 < \xi < x_0 + x$$

Esta es la fórmula de Taylor de segundo orden para la función $y = f(x)$ en x_0.

Una fórmula análoga es la que necesitamos para una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el abierto U de $\mathbb{R}^n$, para poder describir la naturaleza (como posibles extremos locales) de sus puntos críticos. De ella podremos deducir (lo haremos en la próxima sección) un criterio "de la segunda derivada" para identificar extremos locales de la función. Obtener tal fórmula es el objetivo de la presente sección. Por ejemplo, para una función de dos variables $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, definida en el abierto U de $\mathbb{R}^2$, quisiéramos obtener una fórmula que nos diga que en los alrededores del punto $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$ de su dominio, podemos ver a $f(x, y)$ como una función polinomial de grado 2

$$f(x, y) = ax^2 + by^2 + cxy + dx + ey + g$$

(es natural esperar que los coeficientes a, b, c, d, e, g de esta fórmula estén de algún modo comprometidos con las derivadas —parciales— de la función f en el punto $\mathbf{p}$, como ocurre en el caso de una función de una sola variable). Usando luego esta "información cuadrática" para $f(x, y)$ en los alrededores del punto $\mathbf{p}$, podremos establecer (en la siguiente sección) un criterio que nos permita detectar la existencia de extremos locales en los puntos críticos de esta función.

Consideremos pues una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$, y sea $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) \in U$. Supongamos que esta función tiene las derivadas parciales de segundo orden $\frac{\partial^2 f}{\partial x_i \partial x_j}$ continuas en alguna bola B con centro en $\bar{\mathbf{x}}$.

Tomemos la función $g: [-1, 1] \to \mathbb{R}$, definida como

$$g(t) = f(\bar{\mathbf{x}} + t\mathbf{x})$$

donde $\mathbf{x}$ es un punto de $\mathbb{R}^n$ lo suficientemente cerca de $\bar{\mathbf{x}}$ como para que $\bar{\mathbf{x}} \pm \mathbf{x} \in U$. Entonces, g es la composición de la función f con la función $\varphi: [-1, 1] \to \mathbb{R}^n$ dada por $\varphi(t) = \bar{\mathbf{x}} + t\mathbf{x}$, la cual tiene entonces n funciones coordenadas $\varphi_i: [-1, 1] \to \mathbb{R}$ dadas como $\varphi_i(t) = \bar{x}_i + tx_i$, $i = 1, 2, \ldots, n$. Así

$$(f \circ g)(t) = f(\varphi(t)) = f(\bar{\mathbf{x}} + t\mathbf{x}) = g(t), \quad t \in [-1, 1]$$

Apliquemos la fórmula de Taylor de segundo orden a la función g, poniendo en la fórmula (*) $x_0 = 0$, $x = 1$. Nos queda entonces

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(\xi)$$

donde $0 < \xi < 1$

Observamos que $g(1) = f(\bar{\mathbf{x}} + \mathbf{x})$, $g(0) = f(\bar{\mathbf{x}})$. Calculemos $g'(t)$ usando la regla de la cadena

$$g'(t) = \frac{\partial}{\partial t}(f \circ g)(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(t))\varphi_i'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(t))x_i$$

Como $\varphi(0) = \bar{\mathbf{x}}$ nos queda que

$$g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{\mathbf{x}})x_i$$

Calculemos ahora $g''(t)$

$$\begin{aligned} g''(t) &= \frac{\partial}{\partial t}g'(t) = \frac{\partial}{\partial t}\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(t))x_i = \sum_{i=1}^{n} x_i \frac{\partial}{\partial t}\frac{\partial f}{\partial x_i}(\varphi(t)) \\ &= \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\varphi(t))\varphi_j'(t) \right) = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\varphi(t))x_j \right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\varphi(t))x_i x_j \end{aligned}$$

El punto ξ, $0 < \xi < 1$, donde está calculada g'', corresponde al punto $\varphi(\xi) = \bar{\mathbf{x}} + \xi\mathbf{x}$ donde aparecerán calculadas las derivadas parciales de 2° orden de la función f. Nos queda finalmente que, para $\mathbf{x} \in \mathbb{R}^n$ tal que $\bar{\mathbf{x}} + \mathbf{x} \in U$

$$f(\bar{\mathbf{x}} + \mathbf{x}) = \bar{\mathbf{x}} + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{\mathbf{x}})x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{\mathbf{x}} + \xi\mathbf{x})x_i x_j \qquad \text{(T)}$$

donde $0 < \xi < 1$.

Resulta conveniente hacer algunas observaciones con miras a una simplificación de la notación involucrada en esta fórmula. En principio nótese que la expresión

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{\mathbf{x}})x_i$$

se puede ver, tanto como el producto punto del vector grad $f(\bar{\mathbf{x}})$ por el vector $\mathbf{x}$, como el producto de la matriz jacobiana $\mathrm{JF}(\bar{\mathbf{x}})$ por la matriz $\mathbf{x}^t$ (la traspuesta de la matriz de coordenadas de $\mathbf{x}$), o simplemente, también, como la acción de la transformación lineal $f'(\bar{\mathbf{x}})$ (la derivada de f en $\bar{\mathbf{x}}$) en el vector $\mathbf{x}$. Así, esta expresión se puede escribir como grad $f(\bar{\mathbf{x}}) \cdot \mathbf{x}$, $\mathrm{JF}(\bar{\mathbf{x}})\mathbf{x}^t$ o $f'(\bar{\mathbf{x}})\mathbf{x}$, dependiendo de cómo queramos verla.

Lo que resultará interesante será ver de una manera adecuada la expresión

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{\mathbf{x}}+\xi\mathbf{x})x_i x_j$$

Llamemos, por razones de simplificación, $a_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{\mathbf{x}}+\xi\mathbf{x})$ (estos son números reales concretos).

Sea A la matriz $n \times n$ cuyos elementos son a_{ij}, $i, j = 1, 2, \ldots, n$. Identificando el vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ con la matriz $1 \times n$ $\mathbf{x} = [x_1\ x_2 \ldots x_n]$, vemos que

$$\begin{aligned}\mathbf{x}A\mathbf{x}^t &= [x_1\ x_2 \ldots x_n]\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n\end{bmatrix} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j\end{aligned}$$

Como sabemos, a una función $Q: \mathbb{R}^n \to \mathbb{R}$, del tipo $Q(X) = \mathbf{x}A\mathbf{x}^t$, en donde A es una matriz $n \times n$, se le llama *forma cuadrática* (en $\mathbb{R}^n$). A la matriz A se le llama matriz de la forma cuadrática Q (en relación con la base canónica de $\mathbb{R}^n$).

En nuestro caso la matriz A tiene como elementos las derivadas parciales de segundo orden de la función f. Esta matriz recibe un nombre especial.

Definición. Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$ y sea $\bar{\mathbf{x}} \in U$. Suponga que las derivadas parciales de segundo orden $\frac{\partial^2 f}{\partial x_i \partial x_j}$ existen en $\bar{\mathbf{x}}$. A la matriz cuadrada de orden n,

$$A = (a_{ij}) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{\mathbf{x}})\right)_{i,j=1,2,\ldots,n},$$

se le llama MATRIZ HESSIANA (o simplemente HESSIANO) de la función f en $\bar{\mathbf{x}}$ y se denota por $H(\bar{\mathbf{x}})$. ■

Defina, para $\mathbf{x} \in \mathbb{R}^n$ (suficientemente cerca de $\bar{\mathbf{x}}$), la función $r(\mathbf{x})$ como

$$r(\mathbf{x}) = \mathbf{x}(H(\bar{\mathbf{x}}+\xi\mathbf{x}) - H(\bar{\mathbf{x}}))\mathbf{x}^t$$

Entonces la fórmula (T) anterior se puede escribir como

$$f(\bar{\mathbf{x}}+\mathbf{x}) = f(\bar{\mathbf{x}}) + \operatorname{grad} f(\bar{\mathbf{x}}) \cdot \mathbf{x} + \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x})$$

Queremos ver que "el residuo" $r(\mathbf{x})$ tiene la propiedad

$$\lim_{\mathbf{x}\to 0}\frac{r(\mathbf{x})}{\|\mathbf{x}\|^2} = 0$$

En principio téngase presente que si $A = (a_{ij})_{i,j=1,2,\ldots,n}$ es una matriz cuadrada de orden n, entonces

$$\begin{aligned}|\mathbf{x}A\mathbf{x}^t| &= \left|\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j\right| \leq \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}||x_ix_j| \\ &\leq \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|(x_1^2 + x_2^2 + \ldots + x_n^2) = \|\mathbf{x}\|^2\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|\end{aligned}$$

Aplicando este hecho a $r(\mathbf{x})$ nos queda

$$\begin{aligned}|r(\mathbf{x})| &= |\mathbf{x}(H(\bar{\mathbf{x}} + \xi\mathbf{x}) - H(\bar{\mathbf{x}}))\mathbf{x}^t| \\ &\leq \|\mathbf{x}\|^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\frac{\partial^2 f}{\partial x_j\partial x_i}(\bar{\mathbf{x}} + \xi\mathbf{x}) - \frac{\partial^2 f}{\partial x_j\partial x_i}(\bar{\mathbf{x}})\right|\end{aligned}$$

de modo que

$$\frac{|r(\mathbf{x})|}{\|\mathbf{x}\|^2} \leq \sum_{i=1}^{n}\sum_{j=1}^{n}\left|\frac{\partial^2 f}{\partial x_j\partial x_i}(\bar{\mathbf{x}} + \xi\mathbf{x}) - \frac{\partial^2 f}{\partial x_j\partial x_i}(\bar{\mathbf{x}})\right|$$

Recuerde que estamos asumiendo como hipótesis que la función f tiene derivadas parciales de segundo orden continuas en alguna bola B con centro en $\bar{\mathbf{x}}$. Tomando entonces límite en ambas partes de la expresión anterior cuando $\mathbf{x} \to \mathbf{0}$. Por la continuidad de $\frac{\partial^2 f}{\partial x_j\partial x_i}$ vemos que $\frac{\partial^2 f}{\partial x_j\partial x_i}(\bar{\mathbf{x}} + \xi\mathbf{x}) \to \frac{\partial^2 f}{\partial x_j\partial x_i}(\bar{\mathbf{x}})$, por lo que el segundo miembro de esta desigualdad tiende a cero. Esto prueba entonces la propiedad deseada para $r(\mathbf{x})$.

En conclusión, hemos probado la fórmula de Taylor de segundo orden para la función f, resultado que a continuación enunciamos a modo de resumen.

Teorema 4.2.1 (Fórmula de Taylor de segundo orden para funciones de varias variables). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$, tal que las derivadas parciales de segundo orden $\frac{\partial^2 f}{\partial x_j\partial x_i}$ son continuas en alguna bola B con centro en $\bar{\mathbf{x}} \in U$ (por supuesto B contenida en U). Entonces, para todo $\mathbf{x} \in \mathbb{R}^n$ tal que $\bar{\mathbf{x}} + \mathbf{x} \in U$ se tiene la fórmula (de Taylor de segundo orden para la función f en $\bar{\mathbf{x}}$)

$$f(\bar{\mathbf{x}} + \mathbf{x}) = f(\bar{\mathbf{x}}) + \operatorname{grad} f(\bar{\mathbf{x}}) \cdot \mathbf{x} + \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}} + \xi\mathbf{x})\mathbf{x}^t$$

donde $0 < \xi < 1$, la cual se puede escribir también como

$$f(\bar{\mathbf{x}} + \mathbf{x}) = f(\bar{\mathbf{x}}) + \operatorname{grad} f(\bar{\mathbf{x}}) \cdot \mathbf{x} + \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x})$$

en que el residuo $r(\mathbf{x})$ tiene la propiedad

$$\lim_{\mathbf{x}\to\mathbf{0}} \frac{r(\mathbf{x})}{\|\mathbf{x}\|^2} = 0$$

■

Observamos que ya que f es una función con derivadas parciales de segundo orden continuas en una bola B de centro en $\bar{\mathbf{x}}$, cumple las hipótesis del teorema de Schwarz y por lo tanto las derivadas parciales mixtas son iguales. Es decir

$$\frac{\partial^2 f}{\partial x_i \partial y_j}(\bar{\mathbf{x}}) = \frac{\partial^2 f}{\partial x_j \partial y_i}(\bar{\mathbf{x}})$$

Este hecho se puede enunciar de otro modo diciendo que la matriz hessiana $H(\bar{\mathbf{x}})$ es una *matriz simétrica*.

Veamos algunos ejemplos.

Ejemplo 1. Considere la función $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = e^{2x+3y}$. Ciertamente esta función es diferenciable y tiene sus derivadas parciales de todos los órdenes continuas en todo $\mathbb{R}^2$. La fórmula de Taylor de f en el origen se ve como

$$f((0,0) + (x, y)) = f(0,0) + \operatorname{grad} f(0,0) \cdot (x, y) + \frac{1}{2}[x\ y]H(0,0)\begin{bmatrix} x \\ y \end{bmatrix} + r(x, y)$$

donde

$$\lim_{(x,y)\to(0,0)} \frac{r(x, y)}{x^2 + y^2} = 0$$

Hagamos explícita esta fórmula. Tenemos

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2e^{2x+3y}, 3e^{2x+3y})$$

de modo que

$$\operatorname{grad} f(0,0) = (2, 3)$$

Por otra parte

$$H(x, y) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial y \partial x} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 4e^{2x+3y} & 6e^{2x+3y} \\ 6e^{2x+3y} & 9e^{2x+3y} \end{bmatrix}$$

de modo que

$$H(0,0) = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$$

Así, la fórmula de Taylor se ve como

$$f(x, y) = f(0,0) + (2,3) \cdot (x, y) + \frac{1}{2}[x\ y]\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + r(x, y)$$

o sea

$$e^{2x+3y} = 1 + 2x + 3y + 2x^2 + 6xy + \frac{9}{2}y^2 + r(x, y)$$

Insistimos en el signicado de esta fórmula: si quisiéramos ver la función $f(x, y) = e^{2x+3y}$ en los alrededores del origen como un polinomio de grado 2, éste sería $p(x, y) = 1 + 2x + 3y + 2x^2 + 6xy + \frac{9}{2}y^2$, y el error $r(x, y)$ que se estaría cometiendo sería menor que $x^2 + y^2$. Por ejemplo, ponga $x = 0.01$, $y = -0.03$. Se tiene $p(0.01, -0.03) = 0.93245$, en tanto que $f(0.01, -0.03) = 0.9323238$ (el error cometido es de 5.618×10^{-5}, que es menor que $(0.01)^2 + (-0.03)^2 = 1 \times 10^{-3}$). ■

Ejemplo 2. Considere la función $f: \mathbb{R}^3 \to \mathbb{R}$ dada por $u = f(x, y, z) = \operatorname{sen} x + \operatorname{sen} y + \operatorname{sen} z - \operatorname{sen}(x + y + z)$. Esta función es diferenciable y tiene sus derivadas parciales de todos los órdenes continuas en todo $\mathbb{R}^3$. Tomemos el punto $\mathbf{p} = \left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right)$ y escribamos la fórmula de Taylor para f en $\mathbf{p}$.

Se tiene $f\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right) = 4$. Las derivadas parciales de f son

$$\frac{\partial f}{\partial x} = \cos x - \cos(x + y + z), \qquad \frac{\partial f}{\partial y} = \cos y - \cos(x + y + z)$$

$$\frac{\partial f}{\partial z} = \cos z - \cos(x + y + z)$$

en el punto $\mathbf{p}$ se tiene $\frac{\partial f}{\partial x}(\mathbf{p}) = \frac{\partial f}{\partial y}(\mathbf{p}) = \frac{\partial f}{\partial z}(\mathbf{p}) = 0$.

La matriz Hessiana de f es

$$H(x, y, z) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial z \partial x} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial z \partial y} \\ \dfrac{\partial^2 f}{\partial x \partial z} & \dfrac{\partial^2 f}{\partial y \partial z} & \dfrac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

$$= \begin{bmatrix} -\operatorname{sen} x + \operatorname{sen}(x + y + z) & \operatorname{sen}(x + y + z) & \operatorname{sen}(x + y + z) \\ \operatorname{sen}(x + y + z) & -\operatorname{sen} y + \operatorname{sen}(x + y + z) & \operatorname{sen}(x + y + z) \\ \operatorname{sen}(x + y + z) & \operatorname{sen}(x + y + z) & -\operatorname{sen} z + \operatorname{sen}(x + y + z) \end{bmatrix}$$

que en el punto $\mathbf{p}$ queda como

$$H(\mathbf{p}) = \begin{bmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{bmatrix}$$

Entonces la fórmula de Taylor se ve como

$$f\left(\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right) + (x, y, z)\right) = f\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right) + \operatorname{grad} f\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right) \cdot (x, y, z)$$

$$+ \frac{1}{2}[x\ y\ z] H\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} + r(x, y, z)$$

o sea

$$\operatorname{sen}\left(\frac{\pi}{2} + x\right) + \operatorname{sen}\left(\frac{\pi}{2} + y\right) + \operatorname{sen}\left(\frac{\pi}{2} + z\right) - \operatorname{sen}\left(\frac{3\pi}{2} + x + y + z\right)$$

$$= 4 + \frac{1}{2}[x\ y\ z] \begin{bmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + r(x, y, z)$$

$$= 4 - x^2 - y^2 - z^2 - xy - xz - yz + r(x, y, z)$$

en donde

$$\lim_{(x,y,z)\to(0,0,0)} \frac{r(x, y, z)}{x^2 + y^2 + z^2} = 0$$

■

Ejercicios (Capítulo 4, Sección 2)

En los ejercicios 1–9, escriba la fórmula de Taylor de segundo orden para la función dada, en un punto cualquiera $\bar{\mathbf{x}}$ de su dominio, en la forma

$$f(\bar{\mathbf{x}} + \mathbf{x}) = f(\bar{\mathbf{x}}) + \text{grad } f(\bar{\mathbf{x}}) \cdot \mathbf{x} + \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x})$$

Verifique en cada caso la propiedad del residuo

$$\lim_{\mathbf{x}\to\mathbf{0}} \frac{r(\mathbf{x})}{\|\mathbf{x}\|^2} = 0$$

1. $f(x, y) = 5$

2. $f(x, y) = 5x$

3. $f(x, y) = 5x + 8y + 4$

4. $f(x, y) = 10y^2$

5. $f(x, y) = 3xy$

6. $f(x, y) = 2x^2 + 7xy + 5y^2 - 2$

7. $f(x, y) = x^3 + y^3$

8. $f(x, y, z) = 3x - 2y + 15z - 23$

9. $f(x, y, z) = 2x^2 - 5y^2 + 3z^2 + xy - 6xz + 2yz + 1$

10. Escriba la fórmula de Taylor de segundo orden en un punto cualquiera $\mathbf{p} \in \mathbb{R}^n$, para una función lineal $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n + b$.

12. Escriba la fórmula de Taylor de segundo orden en un punto cualquiera $\mathbf{p} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ para la función ("cuadrática") $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$.

13. Constate que la fórmula de Taylor de segundo orden de una función $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ en un punto $\mathbf{p} = (x, y) \in U$, se puede escribir como

$$f(x, y) = f(h, k) + \text{grad } f(h, k) \cdot (x-h,\ y-k) + \frac{1}{2}[x-h\ \ y-k]H(h, k)[x-h\ \ y-k]^t + r(x-h,\ y-k)$$

donde el residuo $r(x - h, y - k)$ tiene la propiedad

$$\lim_{(x,y)\to(h,k)} \frac{r(x - h,\ y - k)}{\|(x - h,\ y - k)\|^2} = 0$$

Véase la fórmula anterior como una manera de "expander" la función $f(x, y)$ como un polinomio en potencias (hasta segundo orden) de $(x - h)$ y $(y - k)$.

14. Después de verificar que una función polinomial del tipo $f(x, y) = Ax^2 + By^2 + Cxy + D$ tiene el residuo de la fórmula de Taylor de segundo orden igual a cero, use el resultado del ejercicio anterior para escribir los polinomios en dos variables dados a continuación en potencias de $(x - h)$ y $(y - k)$, con h y k dados.

a. $3x^2 + 4y^2 - 8xy + 5$, en potencias de $(x - 1)$ y $(y - 2)$

b. $5x^2 - 10y^2 + 14$, en potencias de $(x + 3)$ y $(y - 3)$

c. $x^2 + y^2$, en potencias de $(x - 5)$ y $(y - 12)$

15. Después de verificar que una función polinomial del tipo $f(x, y, z) = Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + G$, tiene el residuo de la fórmula de Taylor de segundo orden igual a cero, tome el resultado del ejercicio 13 para escribir los siguientes polinomios en 3 variables, en potencias de $(x - \alpha)$, $(y - \beta)$ y $(z - \gamma)$, con α, β y γ dados.

a. $x^2 + y^2 + z^2$, en potencias de $(x - 1)$, $(y - 2)$ y $(z - 3)$

b. $2x^2 + 3y^2 - 5z^2 + 3xy + 3$, en potencias de x, $(y + 1)$, $(z - 1)$

c. $x^2 - 2y^2 + 4z^2 + xy + xz - yz + 1$, en potencias de $(x - 1)$, $(y + 1)$, z.

16. Justifique la fórmula aproximada

$$f(x, y) \approx f(h, k) + \text{grad } f(h, k) \cdot (x - h, y - k) + \frac{1}{2}[x - h \ \ y - k]H(h, k)[x - h \ \ y - k]^t$$

para valores de (h, k) cercanos a (x, y). Use este hecho para obtener la siguiente fórmula aproximada para calcular x^y con x y y cerca de 1

$$x^y \approx xy - y + 1$$

Calcule con esta fórmula los valores aproximados de $(0.98)^{0.95}$, $(1.1027)^{0.91}$. Compare con los valores exactos.

17. Obtenga la fórmula aproximada

$$\frac{\cos x}{\cos y} \approx 1 - \frac{1}{2}(x^2 - y^2)$$

para valores de x y y cerca de cero.

18. Obtenga la fórmula aproximada

$$\cos(x + y) \approx 1 - xy - \frac{1}{2}(x^2 + y^2)$$

para valores pequeños de x y y. Utilice este hecho para justificar los siguientes límites

$$\lim_{(x,y)\to(0,0)} \frac{1 - \cos(x + y)}{(x + y)^2} = \frac{1}{2}, \qquad \lim_{(x,y)\to(0,0)} \frac{1 - \cos(x + y)}{(x + y)} = 0$$

19. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U (conteniendo al origen) de $\mathbb{R}^2$, con derivadas parciales de segundo orden continuas en alguna bola B con centro en el

origen. Demuestre que para $(x, y) \in U$ se tiene

$$f(x, y) = f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y + \frac{\partial^2 f}{\partial x \partial y}(0, 0)xy + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(0, 0)x^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(0, 0)y^2 + r(x, y)$$

donde

$$\lim_{(x,y)\to(0,0)} \frac{r(x, y)}{x^2 + y^2} = 0$$

Esta es la fórmula de Taylor de segundo orden para la función f en el origen.

En cada uno de los ejercicios 20—26, obtenga la fórmula de Taylor de segundo orden de la función dada en el origen.

20. $f(x, y) = x^3 + y^3$

21. $f(x, y) = \dfrac{1}{1 + x + y}$

22. $f(x, y) = \dfrac{1}{1 + x^2 + y^2}$

23. $f(x, y) = \ln(1 - x) + \ln(1 - y)$

24. $f(x, y) = e^x \operatorname{sen} y$

25. $f(x, y) = e^x \cos y$

26. $f(x, y) = \arctan(x + y)$

27. Considere la función $z = f(x, y)$ que, en los alrededores del punto $(1, 1, 1)$, está definida implícitamente por

$$z^3 + 3x^2y - y^3z + y^2 - 3x - 1 = 0$$

Obtenga la fórmula de Taylor de segundo orden de tal función en el punto $(1, 1, 1)$. Con esta fórmula, dé estimaciones aproximadas de $f(1.1, 0.9)$ y $f(0.912, 1.087)$.

(*) 28. Considere las funciones $u = u(x, y)$, $v = v(x, y)$ definidas implícitamente, alrededor del punto $(1, 1, 1, 1)$, por

$$F(x, y, u, v) = x + y - u^3 - v^3 = 0$$
$$G(x, y, u, v) = x - y + u^2 - v^2 = 0$$

Obtenga las fórmulas de Taylor de segundo orden de las funciones u y v en el punto $(1, 1)$. Use las fórmulas obtenidas para calcular aproximadamente los valores de las funciones $u = u(x, y)$, $v = v(x, y)$ en los puntos (x, y) indicados a continuación. Para comparar, se proporcionan los valores de estas funciones en tales puntos con una exactitud de 10^{-9}. ¿Qué tan preciso es el cálculo de tales valores con la fórmula de Taylor a medida que nos alejamos del punto $(1, 1)$?

a. $(x, y) = (0.9, 1.1)$. (Valores exactos $u = 1.0476133415$, $v = 0.9473614481$).

b. $(x, y) = (0.7, 1.3)$. (Valores exactos $u = 1.1300288142$, $v = 0.8227789016$).

c. $(x, y) = (0.5, 1.5)$. (Valores exactos $u = 1.1970285668$, $v = 0.6579341834$).

d. $(x, y) = (0.3, 1.7)$. (Valores exactos $u = 1.2469796179$, $v = 0.3936472628$).

4.3 Condiciones suficientes para la existencia de extremos locales

En esta sección se estudiarán resultados generales que aseguran la existencia de extremos locales de funciones diferenciables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ en sus puntos críticos. Esto se logrará con la ayuda de la fórmula de Taylor de segundo orden estudiada en la sección anterior.

Suponga entonces que la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^n$ tiene en $\bar{\mathbf{x}} \in U$ un punto crítico, y que en ese punto satisface las hipótesis del teorema 4.2.1. En tal caso el vector gradiente de f en $\bar{\mathbf{x}}$, constituido por las derivadas parciales de f en $\bar{\mathbf{x}}$, es igual a cero (al vector cero) y por lo tanto la fórmula de Taylor en $\bar{\mathbf{x}}$ se ve como

$$\begin{aligned} f(\bar{\mathbf{x}}+\mathbf{x}) - f(\bar{\mathbf{x}}) &= \operatorname{grad} f(\bar{\mathbf{x}}) \cdot \mathbf{x} + \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x}) \\ &= \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x}) \end{aligned}$$

donde $\text{lím}_{\mathbf{x}\to\mathbf{0}} \frac{r(\mathbf{x})}{\|\mathbf{x}\|^2} = 0$.

Habíamos visto, en la sección 1, que el que la función f tenga o no un extremo local en $\bar{\mathbf{x}}$ depende del comportamiento del signo de $f(\bar{\mathbf{x}}+\mathbf{x}) - f(\bar{\mathbf{x}})$ para $\mathbf{x}$ en una bola B en $\mathbb{R}^n$ con centro en cero. Al ver la fórmula anterior y al recordar que $r(\mathbf{x})$ es un residuo "muy pequeño", parecería que el comportamiento de $f(\bar{\mathbf{x}}+\mathbf{x}) - f(\bar{\mathbf{x}})$ está determinado por el de $\frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t$, con $\mathbf{x}$ pequeño. De hecho así ocurre. Es justamente la matriz hessiana $H(\bar{\mathbf{x}})$ la que tiene la información que necesitamos del signo de $f(\bar{\mathbf{x}}+\mathbf{x}) - f(\bar{\mathbf{x}})$, con $\mathbf{x}$ pequeño. De modo más preciso, se demostrará que si para $\mathbf{x}$ en una bola con centro en $\mathbf{0}$ ($\mathbf{x} \neq 0$) se tiene que $\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t > 0$ ($\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t < 0$), entonces la función f tendrá un mínimo local (máximo local respectivamente) en el punto crítico.

Ahora es cuando usaremos algunos de los resultados estudiados en Algebra Lineal, que se recordaron en la sección 8 del capítulo 1, y que, por consideración al lector, enunciaremos aquí brevemente.

Dada la matriz cuadrada A de orden n, se dice que $\lambda \in \mathbb{R}$ es un valor propio de A si existe un vector $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. Al vector $\mathbf{x}$ se le llama vector propio asociado a λ. El conjunto de vectores $\mathbf{x} \in \mathbb{R}^n$ tales que $A\mathbf{x} = \lambda\mathbf{x}$, unión con el vector cero, forman un espacio vectorial, subespacio de $\mathbb{R}^n$, llamado espacio propio asociado a λ. Este espacio propio es el espacio solución del sistema homogéneo de ecuaciones lineales $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Se sabe que λ es un valor propio de A si y sólo si es una raíz del polinomio característico de A, $p(\lambda) = \det(A - \lambda I)$, el cual es un polinomio de grado n en λ. De aquí se ve que una matriz A de orden n tiene no más de n valores propios (reales). Si la matriz $A = (a_{ij})_{i,j=1,2,\ldots,n}$ es simétrica (es decir, $a_{ij} = a_{ji}$, $i, j = 1, 2, \ldots, n$) entonces A tiene todos sus valores propios reales. Más aún, siendo A simétrica entonces es diagonalizable ortogonalmente, es decir, existe una matriz ortogonal P de orden n (es decir, P es inversible y $P^{-1} = P^t$) que diagonaliza a A, o sea que $P^tAP = D$, donde D es una matriz diagonal que tiene los valores propios de A en su diagonal principal.

La forma cuadrática $Q(\mathbf{x}) = \mathbf{x}A\mathbf{x}^t$, que tiene asociada la matriz A (respecto de la base canónica de $\mathbb{R}^n$) se dice ser:

a. definida positiva, si $Q(\mathbf{x}) > 0\ \forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$,

b. definida negativa, si $Q(\mathbf{x}) < 0\ \forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

Se prueba fácilmente que la forma $Q(\mathbf{x}) = \mathbf{x}A\mathbf{x}^t$ es definida positiva (definida negativa) si y sólo si la matriz A tiene todos sus valores propios positivos (negativos, respectivamente. Ver teorema 1.9.1,

capítulo 1, sección 9). Este es un hecho que usaremos en la discusión que presentamos en esta sección sobre extremos locales (en la demostración del siguiente teorema).

El resultado principal de esta sección está contenido en el siguiente teorema.

Teorema 4.3.1 (Condiciones suficientes para la existencia de extremos locales). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^n$ que tiene en $\bar{\mathbf{x}} \in U$ un punto crítico. Supongamos que en una bola B de $\mathbb{R}^n$ con centro en $\bar{\mathbf{x}}$ las derivadas parciales de f de segundo orden son continuas. Sea $H(\bar{\mathbf{x}})$ la matriz hessiana de f en $\bar{\mathbf{x}}$. Entonces

a. Si la forma cuadrática $Q(\mathbf{x}) = \mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t$ es definida positiva, entonces f tiene un mínimo local en $\bar{\mathbf{x}}$.

b. Si la forma cuadrática $Q(\mathbf{x}) = \mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t$ es definida negativa, entonces f tiene un máximo local en $\bar{\mathbf{x}}$.

Demostración. Como ya se había dicho, nuestro punto de partida es la fórmula de Taylor de la función f en $\bar{\mathbf{x}}$ (las hipótesis hechas sobre f en el teorema, garantizan que podemos usar tal fórmula), la cual se ve, por ser $\bar{\mathbf{x}}$ un punto crítico de f, como

$$f(\bar{\mathbf{x}} + \mathbf{x}) - f(\bar{\mathbf{x}}) = \frac{1}{2}\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x})$$

donde $\lim_{\mathbf{x}\to\mathbf{0}} \frac{r(\mathbf{x})}{\|\mathbf{x}\|^2} = 0$. Demostremos el inciso **a**. Tendremos que demostrar que existe un número $\delta > 0$, radio de una bola B en $\mathbb{R}^n$ con centro en $\mathbf{0}$, tal que $f(\bar{\mathbf{x}} + \mathbf{x}) - f(\bar{\mathbf{x}}) \geq 0$ para toda $\mathbf{x}$ en B (o bien, tal que $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$ para toda $\mathbf{x} \in \mathbb{R}^n$ con $\|\mathbf{x} - \bar{\mathbf{x}}\| < \delta$, lo cual dice que f tiene un mínimo local en $\bar{\mathbf{x}}$). Sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ ($k \leq n$) los valores propios de la matriz $H(\bar{\mathbf{x}})$. Como sabemos, el que la forma $Q(\mathbf{x}) = \mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t$ sea definida positiva es equivalente a que todos los números λ_1, $\lambda_2, \ldots, \lambda_k$ sean positivos. Sea t un número positivo menor que todos los λ_i. Entonces los números $\lambda_i - t$, $i = 1, 2, \ldots, k$ siguen siendo números positivos, los cuales son los valores propios de la matriz $H(\bar{\mathbf{x}}) - tI$ (en efecto, si $\mathbf{x}$ es un vector propio asociado al valor propio λ_i de $H(\bar{\mathbf{x}})$, entonces $(H(\bar{\mathbf{x}}) - tI)\mathbf{x} = H(\bar{\mathbf{x}})\mathbf{x} - t\mathbf{x} = \lambda_i\mathbf{x} - t\mathbf{x} = (\lambda_i - t)\mathbf{x}$, lo que muestra que $\lambda_i - t$ es un valor propio de la matriz $H(\bar{\mathbf{x}}) - tI$). De este modo, que la matriz $H(\bar{\mathbf{x}}) - tI$ tiene todos sus valores propios $\lambda_i - t$ positivos. Por lo tanto la forma cuadrática $Q'(\mathbf{x}) = \mathbf{x}(H(\bar{\mathbf{x}}) - tI)\mathbf{x}^t$ es definida positiva, es decir, $\mathbf{x}(H(\bar{\mathbf{x}}) - tI)\mathbf{x}^t > 0$ para toda $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$, de donde se obtiene que $\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t > t \times \mathbf{x}^t = t\|\mathbf{x}\|^2$, $\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$.

Por otra parte como $\lim_{\mathbf{x}\to\mathbf{0}} \frac{r(\mathbf{x})}{\|\mathbf{x}\|^2} = 0$ (y según la definición de límite —ver capítulo 2, sección 3), dado el número $t > 0$, se da $\delta > 0$ de modo que si $0 < \|\mathbf{x}\| < \delta$ entonces $\left|\frac{r(\mathbf{x})}{\|\mathbf{x}\|^2}\right| < t$; es decir $|r(\mathbf{x})| < t\|\mathbf{x}\|^2$ siempre que $0 < \|\mathbf{x}\| < \delta$. Juntando toda la información, concluímos que $\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t > t\|\mathbf{x}\|^2 > |r(\mathbf{x})|$, o bien que $\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t - |r(\mathbf{x})| > 0$ para toda $\mathbf{x}$ tal que $0 < \|\mathbf{x}\| < \delta$.

Por último, tenemos

$$f(\bar{\mathbf{x}} + \mathbf{x}) - f(\bar{\mathbf{x}}) = \mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t + r(\mathbf{x}) > \mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}^t - |r(\mathbf{x})| > 0$$

para toda $\mathbf{x}$ en una bola B con centro en $\mathbf{0}$ y radio $\delta(\mathbf{x} \neq \mathbf{0})$. Esto es lo que queremos probar.

El inciso **b** se deduce fácilmente del inciso **a**, aplicando el resultado probado a la función $F = -f$. Puesto que $\frac{\partial^2 F}{\partial x_i \partial x_j} = -\frac{\partial^2 f}{\partial x_i \partial x_j}$, la matriz hessiana $H_F(\bar{\mathbf{x}})$ de la función F será la negativa de la matriz hessiana de la matriz $H_f(\bar{\mathbf{x}})$. Sabemos que si λ es un valor propio de la matriz A entonces $-\lambda$ es un valor propio de la matriz $-A$ (prueba: de $A\mathbf{x} = \lambda\mathbf{x}$ se sigue $-A\mathbf{x} = -\lambda\mathbf{x}$). Así pues la matriz

$H_F(\bar{\mathbf{x}})$ tiene todos sus valores propios positivos, y entonces, por el inciso **a**, podemos concluir que $F(\bar{\mathbf{x}}+\mathbf{x})-F(\bar{\mathbf{x}})>0$ o sea $f(\bar{\mathbf{x}}+\mathbf{x})-f(\bar{\mathbf{x}})<0$ para toda $\mathbf{x}$ en una bola B. Esta es la conclusión a la que se quería llegar en este inciso. Q.E.D.

Con las mismas ideas aplicadas en la demostración anterior, es posible construir un argumento que pruebe que (en las mismas condiciones del teorema anterior)

c. Si la matriz $H(\bar{\mathbf{x}})$ tiene (todos) sus valores propios positivos y negativos, entonces la función f tendrá un punto de ensilladura en $\bar{\mathbf{x}}$.

Dejamos como ejercicio para el lector la prueba de este hecho.

El punto clave, entonces, para determinar la naturaleza de un punto crítico $\bar{\mathbf{x}}$, es la forma cuadrática $Q(\mathbf{x})=\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}'$. Más aún, el punto clave es la matriz hessiana $H(\bar{\mathbf{x}})$, pues en ella está contenida la información para concluir que la forma cuadrática es definida positiva o definida negativa. Por ejemplo, investigando los valores propios de $H(\bar{\mathbf{x}})$ podemos concluir:

a. Si todos los valores propios de $H(\bar{\mathbf{x}})$ son positivos, entonces (la forma cuadrática $Q(\mathbf{x})=\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}'$ es definida positiva y) la función f tiene un mínimo local en $\bar{\mathbf{x}}$.

b. Si todos los valores propios de $H(\bar{\mathbf{x}})$ son negativos entonces (la forma cuadrática $Q(\mathbf{x})=\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}'$ es definida negativa y) la función f tiene un máximo local en $\bar{\mathbf{x}}$.

Hay otro criterio muy útil que nos dice cuándo la forma $Q(\mathbf{x})=\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}'$ es definida positiva o definida negativa en términos de los elementos de la matriz $H(\bar{\mathbf{x}})$. Dada una matriz cuadrada $A=(a_{ij})\ i,j=1,\ldots,n$ se consideran las submatrices angulares A_k, $k=1,2,\ldots,n$, definidas como

$$A_1=[a_{11}],\quad A_2=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix},\quad A_3=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{13}&a_{23}&a_{33}\end{bmatrix},\ldots,A_n=A$$

Es decir, la matriz A_k está formada por los elementos de la matriz A que están en su "ángulo superior izquierdo", incluyendo k líneas y k columnas de ella. Se define $\Delta_k=\det A_k$. Se tiene entonces que: la forma cuadrática $Q(\mathbf{x})=\mathbf{x}H(\bar{\mathbf{x}})\mathbf{x}'$ es definida positiva (definida negativa) si y sólo si todos los determinantes $\Delta_1,\Delta_2,\ldots,\Delta_n=\det A$ son números positivos (los determinantes Δ_k tienen signos alternados, $\Delta_1<0$, $\Delta_2>0,\ldots$, respectivamente). Para una demostración de este hecho vea [PiI], capítulo 7, sección 2, teorema 2.3 (pág. 623**). De esta manera tenemos el siguiente nuevo criterio:

a. Si todas las submatrices angulares de la matriz hessiana $H(\bar{\mathbf{x}})$ tienen determinantes positivos, entonces la función f tiene un mínimo local en $\bar{\mathbf{x}}$.

b. Si las submatrices angulares de la matriz hessiana $H(\bar{\mathbf{x}})$ tienen determinantes de signo alternado (comenzando con un valor negativo, o sea $\frac{\partial^n f}{\partial x_1^n}(\bar{\mathbf{x}})<0$), entonces la función f tiene un máximo local en $\bar{\mathbf{x}}$.

Veamos algunos ejemplos.

Ejemplo 1. Sea $f:\mathbb{R}^2\to\mathbb{R}$ la función dada por

$$f(x,y)=2(x-1)^2+3(y-2)^2$$

Los puntos críticos de f se obtienen de

$$\frac{\partial f}{\partial x}=4(x-1)=0\qquad\frac{\partial f}{\partial y}=6(y-2)=0$$

sistema que tiene por solución a $x = 1$, $y = 2$. Entonces $(1, 2)$ es el único punto crítico de la función. La matriz hessiana de f es

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$$

que en $(1, 2)$ es ella misma. Se ve que los valores propios de esta matriz son 4 y 6 (ambos positivos) o bien, que los determinantes de las submatrices angulares son 4 y 24 (también ambos positivos) por lo que concluimos que la función f tiene en el punto $(1, 2)$ un mínimo local (la gráfica de la función es un paraboloide elíptico que tiene su vértice en $(1, 2)$ y abre hacia arriba). ■

Ejemplo 2. Consideremos la función $f: \mathbb{R}^3 \to \mathbb{R}$

$$f(x, y, z) = \operatorname{sen} x + \operatorname{sen} y + \operatorname{sen} z - \operatorname{sen}(x + y + z)$$

(ver ejemplo 2 de la sección anterior). El punto $\mathbf{p} = \left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right)$ es un punto crítico de f. En ese punto la matriz hessiana de f es

$$H(\mathbf{p}) = \begin{bmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{bmatrix}$$

Los determinantes de las submatrices angulares son

$$\Delta_1 = \det[-2] = -2, \quad \Delta_2 = \det \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix} = 3, \quad \Delta_3 = \det H(\mathbf{p}) = -4$$

Puesto que son de signos alternados (con $\Delta_1 < 0$), concluimos que la función f tiene en $\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right)$ un máximo local. (Conclusión a la que se hubiera podido llegar viendo que el polinomio característico $H(\mathbf{p})$ es $\det(H(\mathbf{p}) - \lambda I) = -\lambda^3 - 6\lambda^2 - 9\lambda - 4 = -(\lambda + 1)^2(\lambda + 4)$ y que por lo tanto los valores propios de $H(\mathbf{p})$ son $\lambda = -1$ y $\lambda = -4$, ambos negativos). Este máximo local vale $f\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right) = 4$. ■

Ejemplo 3. Sea $f: \mathbb{R}^4 - \{(0, 0, 0, 0)\} \to \mathbb{R}$ dada por

$$f(x, y, z, u) = x + \frac{y}{x} + \frac{z}{y} + \frac{u}{z} + \frac{1}{u}$$

Los puntos críticos de f se obtienen de

$$\frac{\partial f}{\partial x} = 1 - \frac{y}{x^2} = 0 \qquad \frac{\partial f}{\partial y} = \frac{1}{x} - \frac{z}{y^2} = 0$$

$$\frac{\partial f}{\partial z} = \frac{1}{y} - \frac{u}{z^2} = 0 \qquad \frac{\partial f}{\partial u} = \frac{1}{z} - \frac{1}{u^2} = 0$$

Este sistema se satisface para $x = y = z = u = 1$. Obtengamos la matriz hessiana de f en $\mathbf{p} = (1, 1, 1, 1)$.

$$H(x, y, z, u) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} & \dfrac{\partial^2 f}{\partial x \partial u} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} & \dfrac{\partial^2 f}{\partial y \partial u} \\ \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} & \dfrac{\partial^2 f}{\partial z \partial u} \\ \dfrac{\partial^2 f}{\partial u \partial x} & \dfrac{\partial^2 f}{\partial u \partial y} & \dfrac{\partial^2 f}{\partial u \partial z} & \dfrac{\partial^2 f}{\partial u^2} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{2y}{x^3} & -\dfrac{1}{x^2} & 0 & 0 \\ -\dfrac{1}{x^2} & \dfrac{2z}{y^3} & -\dfrac{1}{y^2} & 0 \\ 0 & -\dfrac{1}{y^2} & \dfrac{2u}{z^3} & -\dfrac{1}{z^2} \\ 0 & 0 & -\dfrac{1}{z^2} & \dfrac{2}{u^3} \end{bmatrix}$$

de modo que

$$H(1, 1, 1, 1) = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

Las submatrices angulares de esta matriz tienen por determinantes

$$\Delta_1 = \det[2] = 2, \quad \Delta_2 = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 3$$

$$\Delta_3 = \det \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 2, \quad \Delta_4 = \det \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} = 5$$

Como son positivos, concluimos entonces que la función f tiene un mínimo local en $\mathbf{p}$, que vale $f(1, 1, 1, 1) = 5$. ■

Ejemplo 4. Sea $f: \subseteq \mathbb{R}^3 \to \mathbb{R}$ la función dada por

$$f(x, y, z) = e^{-x^2} + e^{-y^2} + z^2$$

Las derivadas parciales de f son

$$\frac{\partial f}{\partial x} = -2xe^{-x^2}, \quad \frac{\partial f}{\partial y} = -2ye^{-y^2}, \quad \frac{\partial f}{\partial z} = 2z$$

de donde se ve que el único punto crítico es (0, 0, 0). La matriz hessiana es

$$H(x, y, z) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

$$= \begin{bmatrix} (4x^2 - 2)e^{-x^2} & 0 & 0 \\ 0 & (4y^2 - 2)e^{-y^2} & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

que en el origen se ve como

$$H(0, 0, 0) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Los valores propios de la matriz son $\lambda_1 = -2$ y $\lambda_2 = 2$, por lo que, según la observación hecha después de la demostración del teorema 4.3.1, la función f tiene un punto de ensilladura en el origen. Advertimos que, en este caso, decir que la función $f(x, y, z) = e^{-x^2} + e^{-y^2} + z^2$ tiene un punto de ensilladura en el origen nada tiene que ver con "la forma de la gráfica de f cerca del origen". El nombre "punto de ensilladura" se usa por extensión para funciones de n variables, aunque está inspirado en el caso concreto $n = 2$ (en la gráfica de $f(x, y) = x^2 - y^2$ cerca del origen). Lo que nos dice tal calificativo para el origen es que $f(x, y, z) - f(0, 0, 0) = e^{-x^2} + e^{-y^2} + z^2 - 2$ tiene signos positivos y negativos en cualquier bola B de $\mathbb{R}^3$ con centro en el origen. De hecho, obsérvese que para cualquier r, los puntos de la forma $(r, r, 0)$ son tales en los que $f(x, y, z) - f(0, 0, 0) = 2e^{-r^2} - 2 < 0$, en tanto que para puntos de la forma $(0, 0, r)$ se tiene $f(x, y, z) - f(0, 0, 0) = 2 + r^2 - 2 = r^2 > 0$. ■

En los criterios desarrollados en esta sección nada se ha dicho del caso en que la matriz hessiana $H(\mathbf{p})$, $\mathbf{p}$ un punto crítico de la función f, tenga algún valor propio igual a cero. Acontece que con este caso los criterios aquí estudiados no dicen nada acerca de la naturaleza del punto crítico $\mathbf{p}$. Se podrían establecer resultados para situaciones como ésta usando la fórmula de Taylor de orden 3 de la función f en $\mathbf{p}$. Esto, por supuesto, resulta mucho más complicado que lo aquí visto y no se expondrá en el texto. Veamos, sin embargo, un ejemplo concreto que ilustra esta situación.

Ejemplo 5. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función dada por

$$f(x, y) = ax^2 + by^4$$

Tenemos

$$\frac{\partial f}{\partial x} = 2ax, \quad \frac{\partial f}{\partial y} = 4by^3$$

de donde se ve que el origen (0, 0) es el único punto crítico de la función. La matriz hessiana es

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2a & 0 \\ 0 & 12by^2 \end{bmatrix}$$

que en el origen se ve como

$$H(0,0) = \begin{bmatrix} 2a & 0 \\ 0 & 0 \end{bmatrix}$$

Se tiene entonces que $\lambda = 0$ es un valor propio de la matriz $H(0,0)$. Como dijimos nada se puede concluir en este caso de la naturaleza del punto crítico $(0,0)$ para la función f. De hecho, puede ocurrir cualquier cosa en ese punto: si $a = b = 1$, la función $f(x,y) = x^2 + y^4$ tiene un mínimo local en $(0,0)$, pues es claro que para una bola (cualquiera) de $\mathbb{R}^2$ de centro en $(0,0)$ se tiene $f(x,y) = x^2 + y^4 \geq 0 = f(0,0)$, $\forall (x,y) \in B$; si $a = b = -1$, la función $f(x,y) = -x^2 - y^4$ tiene un máximo local en el origen, pues $f(x,y) = -x^2 - y^4 \leq 0 = f(0,0)$, $\forall (x,y) \in B$. Si $a = -1$ y $b = 1$, la función $f(x,y) = -x^2 + y^4$ tiene un punto de ensilladura en el origen, pues en cualquier bola B de $\mathbb{R}^2$ con centro en el origen, digamos $B = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 < \epsilon^2\}$ se tiene puntos $(r,0)$ y $(0,s)$ de modo que $f(r,0) = -r^2 < 0 = f(0,0)$ (con $r^2 < \epsilon^2$) y $f(0,s) = s^4 > 0 = f(0,0)$ (con $s^2 < \epsilon^2$). ■

Ejercicios (Capítulo 4, Sección 3)

En los ejercicios 1–5, suponga que la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^2$, tiene en $\mathbf{p} = (x_0, y_0)$ un punto crítico, y tiene derivadas parciales de segundo orden continuas en una bola B con centro en $\mathbf{p}$. Dada la matriz hessiana de f en $\mathbf{p}$, determine (cuando concluya) si se trata de un máximo local, mínimo local o un punto de ensilladura.

1. $Hf(\mathbf{p}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

2. $Hf(\mathbf{p}) = \begin{bmatrix} -2 & 5 \\ 5 & 4 \end{bmatrix}$

3. $Hf(\mathbf{p}) = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

4. $Hf(\mathbf{p}) = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$

5. $Hf(\mathbf{p}) = \begin{bmatrix} -9 & 10 \\ 10 & -16 \end{bmatrix}$

En los ejercicios 6–10, suponga que la función $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ definida en el conjunto abierto U de $\mathbb{R}^3$, tiene en $\mathbf{p} = (x_0, y_0, z_0)$ un punto crítico, y tiene derivadas parciales de segundo orden continuas en una bola B con centro en $\mathbf{p}$. Dada la matriz hessiana de f en $\mathbf{p}$, determine (cuando concluya) si se trata de un máximo local, mínimo local o un punto de ensilladura.

6. $Hf(\mathbf{p}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

7. $Hf(\mathbf{p}) = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 3 & 4 \\ 4 & 4 & 4 \end{bmatrix}$

8. $Hf(\mathbf{p}) = \begin{bmatrix} 4 & 5 & 10 \\ 5 & 3 & -4 \\ 10 & -4 & 7 \end{bmatrix}$

9. $Hf(\mathbf{p}) = \begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

10. $Hf(\mathbf{p}) = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$

11. Dé un ejemplo concreto de una función $f: \mathbb{R}^2 \to \mathbb{R}$ y un punto $\mathbf{p} = (x_0, y_0)$, tal que $\mathbf{p}$ sea un punto crítico de f, y que la matriz hessiana de f en $\mathbf{p}$ sea la matriz del ejercicio: **a.** 1; **b.** 4.

12. Dé un ejemplo concreto de una función $f: \mathbb{R}^3 \to \mathbb{R}$ y un punto $\mathbf{p} = (x_0, y_0, z_0)$, tal que $\mathbf{p}$ sea un punto crítico de f, y que la matriz hessiana de f en $\mathbf{p}$ sea la matriz del ejercicio: **a.** 6; **b.** 9

13. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = ax^2 + by^2$. Habiendo verificado que el origen es un punto crítico de f, clasifique todas las posibilidades que tiene este punto de ser extremo local o punto de ensilladura de f, en términos de los coeficientes a y b.

14. Sea $f: \mathbb{R}^3 \to \mathbb{R}$ la función $f(x, y, z) = ax^2 + by^2 + cz^2$. Habiendo verificado que el origen es un punto crítico de f, clasifique todas las posibilidades que tiene este punto de ser extremo local o punto de ensilladura de f, en términos de los coeficientes a, b y c.

En los ejercicios 15–25, la función f tiene un punto crítico en el origen, donde vale 0. Suponga que f tiene derivadas parciales de segundo orden continuas en una bola con centro en el origen. Se da la fórmula de Taylor de segundo orden de la función f en el punto crítico. Determine en cada caso (cuando esto sea posible) la naturaleza del punto crítico $\mathbf{p}$.

15. $f(x, y) = 2x^2 + 8y^2 - 3xy + r(x, y)$

16. $f(x, y) = 3x^2 + 5y^2 + xy + r(x, y)$

17. $f(x, y) = x^2 + y^2 - xy + r(x, y)$

18. $f(x, y) = -3x^2 - 5y^2 + 2xy + r(x, y)$

19. $f(x, y) = -3x^2 - 5y^2 + 7xy + r(x, y)$

20. $f(x, y, z) = 17x^2 + 14y^2 + 14z^2 - 4xy - 4xz - 8yz + r(x, y, z)$

21. $f(x, y, z) = -6x^2 - 5y^2 - 7z^2 + 3xy - 3xz + r(x, y, z)$

22. $f(x, y, z) = -x^2 - 2y^2 - 3z^2 + r(x, y, z)$

23. $f(x, y, z) = x^2 - y^2 + z^2 + r(x, y, z)$

24. $f(x, y, z) = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz + r(x, y, z)$

25. $f(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + r(x_1, x_2, x_3, x_4)$

En los ejercicios 26–34, determine (si los hay) los extremos locales y/o puntos de ensilladura de las funciones dadas de dos variables.

26. $f(x, y) = x^2 + 4y^2 - 3xy + x + y - 1$

27. $f(x, y) = x^2 - y^2 - xy$

28. $f(x, y) = 2x^2 + 2y^2 + 12xy + 3x + 4y - 2$

29. $f(x, y) = x^3y^3 + 3x + 3y + 1$

30. $f(x, y) = x^3 - y^3 - 3x + 3y - 1$

31. $f(x, y) = x^4 - y^2 - x^2 + 5y - 2$

32. $f(x, y) = x \ln y + x$

33. $f(x, y) = \arctan(x^2) - \arctan(y^2)$

34. $f(x, y) = \ln(1 + \exp(-x^2) + \exp(-y^2))$

En los ejercicios 35–56, determine (si los hay) los extremos locales y/o puntos de ensilladura de las funciones dadas de tres variables.

35. $f(x, y, z) = x^4 + y^4 + z^4 + 4x + 4y + 32z + 1$

36. $f(x, y, z) = -2x^4 - y^4 - z^4 + 8x + 4y + 4z - 2$

37. $f(x, y, z) = x^4 + 2y^4 - z^4 + 10x + 12y + 9z$

38. $f(x, y, z) = \frac{1}{3}x^3 - x + 2 - y^2 + 2y - z^2 + 2z$

39. $f(x, y, z) = \frac{1}{3}x^3 + x + y^2 + 2y - z^2 + 2z + 3$

40. $f(x, y, z) = x^3 + 3x^2 - 2y^2 + 4y - 2z^2 + 6z + 2$

41. $f(x, y, z) = 2(x^3 + x^2) + 3x + y^2 - 10y + z + 12$

42. $f(x, y, z) = 15x^3 + 6x^2 - x + 2y^2 + y + 5z^2 + 10z - 2$

43. $f(x, y, z) = x^3 + y^2 + xy + xz + yz + 4y + z - 3$

44. $f(x, y, z) = 2x^3 + 2y^2 + z^2 + 2xy + xz + yz + z - 1$

45. $f(x, y, z) = 2x^3 + 2y^2 + z^2 + 2xy + 3yz$

46. $f(x, y, z) = x^2 + y^2 + z^2 + xy + xz + yz + x + y + z + 2$

47. $f(x, y, z) = -2x^2 - y^2 - 3z^2 + xy + 2x + 2y + 3z$

48. $f(x, y, z) = 2x^2 + 3y^2 + 2z^2 + 2xy + xz + yz + 2x + 4y + 6z$

49. $f(x, y, z) = 6x^2 + 5y^2 + 4z^2 + 3xy + 2xz + yz$

50. $f(x, y, z) = x^2 + 2y^2 + z^2 + 2xy + xz + 2yz + x + 2y + z + 2$

51. $f(x, y, z) = -x^2 - 2y^2 - z^2 - 2xy - xz - 2yz - x - 2y - z - 2$

52. $f(x, y, z) = x^2 + 2y^2 + 3z^2 + xy + 2xz + 3yz + x + 2y + 3z + 1$

53. $f(x, y, z) = -x^2 - y^2 - z^2 + xz + y$

54. $f(x, y, z) = -2x^2 - 3y^2 - z^2 + xy + yz + x + y + z + 2$

55. $f(x, y, z) = 10x^2 - (y^2 + z^2 + xy + xz + yz + x + y + z + 1)$

56. $f(x, y, z) = x^2 + 3y^2 + 5z^2 + 7xy + 9xz + 11yz + 13x + 15y + 17z + 19$

57. Sea $g: \mathbb{R} \to \mathbb{R}$ una función diferenciable. Suponga que g tiene solamente una raíz en el punto x_0 y que $g'(x_0) > 0$. Estudie la naturaleza de los puntos críticos de las siguientes funciones $f: \mathbb{R}^2 \to \mathbb{R}$

a. $f(x, y) = \int_x^y g(t)\, dt$

b. $f(x, y) = \int_{-x}^y g(t)\, dt$

c. $f(x, y) = \int_x^{-y} g(t)\, dt$

d. $f(x, y) = \int_{-x}^{-y} g(t)\, dt$

58. Sea $g: \mathbb{R} \to \mathbb{R}$ una función diferenciable. Suponga que la gráfica de g cruza al eje x solamente en el origen de coordenadas. Estudie la naturaleza de los puntos críticos de la función $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \int_0^{x-1} g(t)dt + \int_0^{y-1} g(t)dt$$

en cada uno de los siguientes casos: **a.** $g'(0) > 0$, **b.** $g'(0) < 0$.

59. Sea $g: \mathbb{R} \to \mathbb{R}$ una función diferenciable. Suponga que esta función no tiene raíces. Determine la naturaleza de los puntos críticos de la función $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \int_0^{(x-1)^2} g(t)dt + \int_0^{(y-1)^2} g(t)dt$$

en cada uno de los siguientes casos: **a.** $g(0) > 0$, **b.** $g(0) < 0$

60. Sea $g: \mathbb{R} \to \mathbb{R}$ una función diferenciable tal que $g(1) = g(2)$. Considere la función $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \int_{xy}^{x+y} g(t)dt$$

Demuestre que f tiene un punto crítico en $(1, 1)$. Estudie la naturaleza de este punto crítico en cada uno de los siguientes casos:

a. $g(1) = 0$, $g'(1) = 1$, $g'(2) = 2$

b. $g(1) = 3$, $g'(1) = 3$, $g'(2) = 4$

61. Sea $g: \mathbb{R} \to \mathbb{R}$ una función diferenciable. Suponga que la gráfica de g pasa por el origen. Demuestre que la función $f: \mathbb{R}^3 \to \mathbb{R}$

$$f(x, y, z) = z^2 + \int_{-x}^{y} g(t)dt$$

tiene un punto crítico en $(0, 0, 0)$. Determine la naturaleza de este punto crítico suponiendo que $g'(0) \neq 0$.

62. Considere la función $f(x, y) = \ln x \ln y$. Compruebe que el punto $(1, 1)$ es el único punto crítico de f y que el criterio desarrollado en esta sección no permite determinar la naturaleza de este punto (como posible extremo local). Dé argumentos que muestren que, de hecho, la función f tiene en $(1, 1)$ un punto de ensilladura.

63. Demuestre que la función $f(x_1, x_2, \ldots, x_n) = (\ln x_1)(\ln x_2) \cdots (\ln x_n)$ tiene un punto crítico en $(1, 1, \ldots, 1)$ el cual es un punto de ensilladura.

Los ejercicios 64–69, en los que se pide estudiar la naturaleza de los puntos críticos de las funciones indicadas, requieren la utilización de técnicas numéricas para localizar tales puntos, como el método de Newton estudiado en la sección 7 del capítulo 3.

64. Determine la naturaleza de los puntos críticos de la función $f(x, y) = x^4 + y^4 - 3x^3y - 2xy^2 + xy + 3x - 3y$.

65. Determine la naturaleza de los puntos críticos de la función $f(x, y) = \ln x \ln y - 0.25x - 0.4y$.

66. Determine la naturaleza de los puntos críticos de la función $f(x, y) = (1 + y^2)\ln\cosh x + 2x^2 + 3y^2 - x - y + 1$. (Ver ejercicio 3 de la sección 7 del capítulo 3, y ejercicio 45 de la sección 1 de este capítulo).

67. Determine la naturaleza de los puntos críticos de la función $f(x, y) = y \operatorname{sen}(x + y) + x\cos(x - y)$, que se encuentran dentro del cuadrado $[-3, 3] \times [-3, 3]$.

68. Encontrar 10 puntos críticos de la función $f(x, y) = y\ln(2 + \cos x) + x \operatorname{sen} y$, en el cuadrado $[1, 11] \times [1, 11]$. Determine la naturaleza de estos puntos críticos.

69. Determine la naturaleza de los puntos críticos de la función $f(x, y, z) = (x^3 + y^2 + z)\exp(-x^2 - y^2 - z^2) - xy - xz - yz$.

4.4 Casos de dos variables. Ejemplos

En esta sección sólo vamos a particularizar en el (importante) caso $n = 2$ los resultados obtenidos en la sección anterior y también presentaremos algunos ejemplos "prácticos" en cuya solución interviene la determinación de extremos locales de funciones de dos variables.

Teorema 4.4.1 (Criterio de la segunda derivada para la determinación de extremos locales para funciones de dos variables). Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el conjunto abierto U de $\mathbb{R}^2$ tal que en una bola B con centro en el punto crítico $(x_0, y_0) \in U$ sus derivadas parciales de segundo orden son continuas. Sea

$$A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0), \quad B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0), \quad C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$$

a. Si $B^2 - AC < 0$ y $A > 0$, entonces la función f tiene un mínimo relativo en (x_0, y_0).

b. Si $B^2 - AC < 0$ y $A < 0$, entonces la función f tiene un máximo relativo en (x_0, y_0).

c. Si $B^2 - AC > 0$, entonces la función f tiene un punto de ensilladura en (x_0, y_0).

d. Si $B^2 - AC = 0$, no se puede afirmar nada acerca de la naturaleza del punto crítico (x_0, y_0). ∎

Demostración. En este caso la matriz hessiana de f en (x_0, y_0) es

$$H(x_0, y_0) = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Los casos **a** y **b** se siguen fácilmente del criterio de los signos de los determinantes de las submatrices angulares de $H(x_0, y_0)$. Sin embargo, presentamos un argumento con valores propios de $H(x_0, y_0)$ con el que también podemos demostrar **c** y **d**. El polinomio característico de $H(x_0, y_0)$ es

$$p(\lambda) = \lambda^2 - (A + C)\lambda + AC - B^2$$

Sean λ_1 y λ_2 las raíces de $p(\lambda)$ (es decir, los valores propios de $H(x_0, y_0)$). Sabemos que

$$\lambda_1 + \lambda_2 = A + C, \quad \lambda_1\lambda_2 = AC - B^2$$

Si $B^2 - AC < 0$ (caso **a** y **b**), entonces $\lambda_1\lambda_2 > 0$ por lo que λ_1 y λ_2 tienen el mismo signo. También A y C deben tener el mismo signo (caso contrario se tendría $B^2 - AC \geq 0$). Entonces, como $\lambda_1 + \lambda_2 = A + C$ concluimos que los signos de λ_1, λ_2, A y C deben ser iguales. Entonces, si $A > 0$, se tiene que λ_1 y λ_2 deben ser positivos y por lo tanto (la matriz hessiana tiene sus dos valores propios positivos) existe un mínimo local en (x_0, y_0). Esto prueba el inciso **a**. Del mismo modo, si $A < 0$, los dos valores propios de $H(x_0, y_0)$, λ_1 y λ_2, son negativos, y por lo tanto existe un máximo local en (x_0, y_0). Esto prueba **b**. Si $B^2 - AC > 0$ entonces $\lambda_1\lambda_2 < 0$ y así, los valores propios de $H(x_0, y_0)$ tienen signos distintos, de lo cual se concluye que existe un punto de ensilladura en (x_0, y_0). Esto prueba **c**. Por último, si $B^2 - AC = 0$, se tiene $\lambda_1\lambda_2 = 0$, de donde uno de los valores propios de $H(x_0, y_0)$ debe ser igual a cero, y, como ya se había dicho en la sección anterior, en este caso no se puede concluir nada con este criterio. Q.E.D.

Ejemplo 1. Considere la función $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = 2x^4 + y^4 - 4x^2 - 2y^2$$

Obtengamos sus puntos críticos

$$\frac{\partial f}{\partial x} = 8x^3 - 8x = 0 \qquad \frac{\partial f}{\partial y} = 4y^3 - 4y = 0$$

Se ve que existen 9 puntos críticos, a saber $\mathbf{p}_1 = (0, 0)$, $\mathbf{p}_2 = (0, 1)$, $\mathbf{p}_3 = (0, -1)$, $\mathbf{p}_4 = (1, 0)$, $\mathbf{p}_5 = (1, 1)$, $\mathbf{p}_6 = (1, -1)$, $\mathbf{p}_7 = (-1, 0)$, $\mathbf{p}_8 = (-1, 1)$, $\mathbf{p}_9 = (-1, -1)$. Se tiene

$$\frac{\partial^2 f}{\partial x^2} = 24x^2 - 8, \quad \frac{\partial^2 f}{\partial x \partial y} = 0, \quad \frac{\partial^2 f}{\partial y^2} = 12y^2 - 4$$

La siguiente tabla resume los resultados a la luz del teorema anterior

Punto (x_0, y_0)	$A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)$	$B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$	$C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$	$B^2 - AC$	Conclusión
$(0, 0)$	-8	0	-4	-32	Máximo local $f(0, 0) = 0$
$(0, 1)$	-8	0	$+8$	$+64$	Punto de ensilladura $f(0, 1) = -1$
$(0, -1)$	-8	0	$+8$	$+64$	Punto de ensilladura $f(0, -1) = -1$
$(1, 0)$	$+16$	0	-4	$+64$	Punto de ensilladura $f(1, 0) = -2$
$(1, 1)$	$+16$	0	$+8$	-128	Mínimo local $f(1, 1) = -3$
$(1, -1)$	$+16$	0	$+8$	-128	Mínimo local $f(1, -1) = -3$
$(-1, 0)$	$+16$	0	-4	$+64$	Punto de ensilladura $f(-1, 0) = -2$
$(-1, 1)$	$+16$	0	$+8$	-128	Mínimo local $f(-1, 1) = -3$
$(-1, -1)$	$+16$	0	$+8$	-128	Mínimo local $f(-1, -1) = -3$

■

Ejemplo 2. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ la función $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$. Escribamos $r = x^2 + y^2$ de modo que la función considerada se ve como $f(r) = re^{-r}$. Obtengamos los puntos críticos de f

$$\frac{\partial f}{\partial x} = f'(r)\frac{\partial r}{\partial x} = (1 - r)e^{-r}2x = 0 \qquad \frac{\partial f}{\partial y} = f'(r)\frac{\partial r}{\partial y} = (1 - r)e^{-r}2y = 0$$

Estas ecuaciones se satisfacen con $x = y = 0$, o bien con $r = 1$, y la última condición se satisface por todos los puntos del círculo unitario $x^2 + y^2 = 1$. Tenemos así que esta función f posee una infinidad de puntos críticos. Obtengamos las derivadas parciales de segundo orden de f

$$\frac{\partial^2 f}{\partial x^2} = 2e^{-r}[1 - r + 4x^2(r - 2)] \quad \frac{\partial^2 f}{\partial x \partial y} = 4xy(r - 2)e^{-r} \quad \frac{\partial^2 f}{\partial y^2} = 2e^{-r}[1 - 2 + 4y^2(r - 2)]$$

Para el punto crítico $(0, 0)$ se tiene $r = 0$ y

$$A = \frac{\partial^2 f}{\partial x^2}(0, 0) = 2, \quad B = \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 0, \quad C = \frac{\partial^2 f}{\partial y^2}(0, 0) = 2$$

y entonces $B^2 - AC = -4 < 0$ y $A = 2 > 0$, por lo que la función f tiene un mínimo local en $(0, 0)$ que vale $f(0, 0) = 0$. Por otra parte, para los puntos $r = 1$ se tiene

$$A = \left.\frac{\partial^2 f}{\partial x^2}\right|_{r=1} = -8x^2e^{-1}, \qquad B = \left.\frac{\partial^2 f}{\partial x \partial y}\right|_{r=1} = -4xye^{-1}, \qquad C = \left.\frac{\partial^2 f}{\partial y^2}\right|_{r=1} = -8y^2e^{-1}$$

de donde $B^2 - AC = 16x^2y^2e^{-2} - 64x^2y^2e^{-2} = -48x^2y^2e^{-2} < 0$ y $A = -8x^2e^{-1} < 0$. En realidad esta condición se da si y sólo si $x \neq 0$. Es decir, en los puntos $(0, \pm 1)$ no podremos concluir nada con este criterio. Sin embargo, aplicando un argumento de continuidad de la función f, incluimos también estos puntos en nuestras conclusiones. Resulta entonces que en todos los puntos en que $r = x^2 + y^2 = 1$ existen máximos locales de la función f, que valen $f(x, y)\big|_{r=1} = e^{-1}$.

Resulta que nuestra superficie $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$ es en realidad una superficie de revolución que se obtiene al girar alrededor del eje y la curva $y = x^2e^{-x^2}$, la cual tiene el aspecto de la (figura 1).

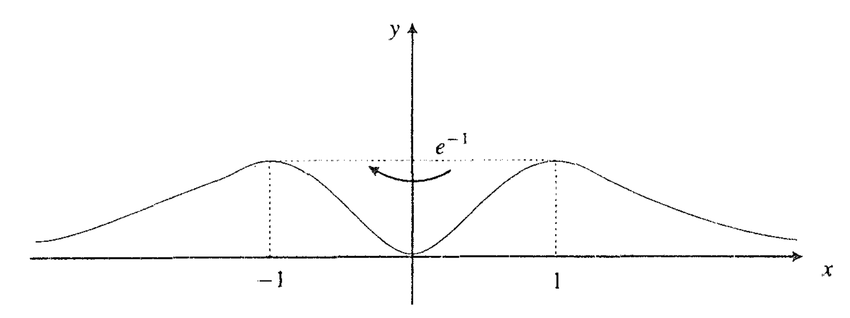


Figura 1. Gráfica de la función $y = x^2e^{-x^2}$.

Siendo entonces este el aspecto de la superficie $z = (x^2 + y^2)e^{-(x^2+y^2)}$

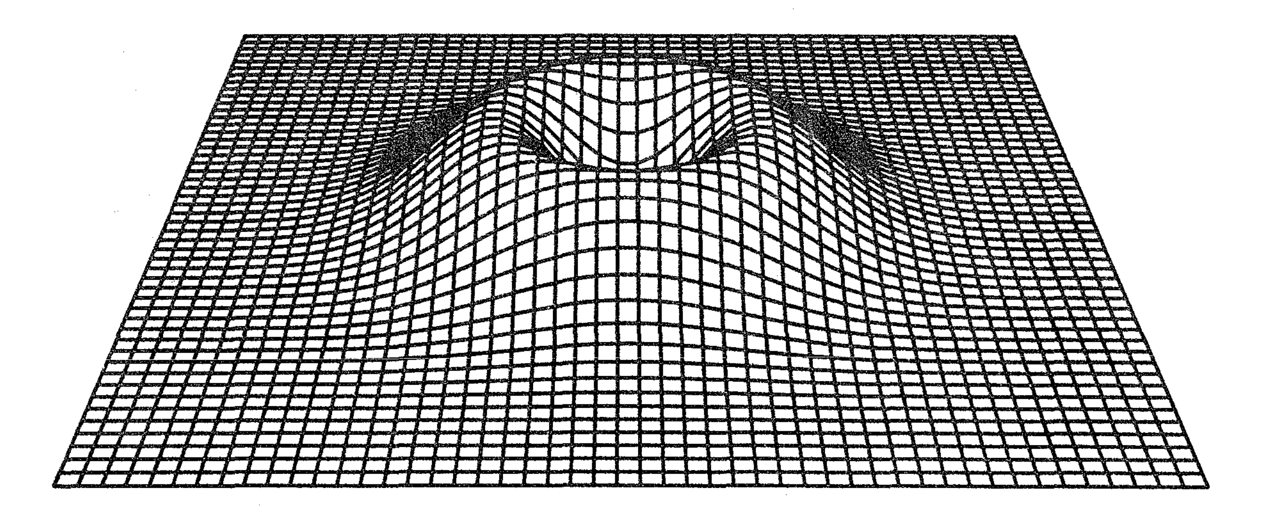


Figura 2. La superficie $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$. ∎

Ejemplo 3. Sea $f: \mathbb{R}^2 - \{0, 0\} \to \mathbb{R}$ la función

$$f(x, y) = 2x - 3y + \frac{1}{2}\ln(x^2 + y^2) + 5\arctan\frac{y}{x}$$

Investiguemos sus puntos críticos

$$\frac{\partial f}{\partial x} = 2 + \frac{x}{x^2+y^2} + 5\frac{1}{1+\frac{y^2}{x^2}}\left(-\frac{y}{x^2}\right) = \frac{2(x^2+y^2)+x-5y}{x^2+y^2} = 0$$

$$\frac{\partial f}{\partial y} = -3 + \frac{y}{x^2+y^2} + 5\frac{1}{1+\frac{y^2}{x^2}}\left(\frac{1}{x^2}\right) = \frac{-3(x^2+y^2)+y+5}{x^2+y^2} = 0.$$

Resolviendo este sistema encontramos el único punto crítico (1, 1). Las derivadas parciales de 2° orden de f en (1, 1) son

$$A = \frac{\partial^2 f}{\partial x^2}(1,1) = \frac{5}{2}, \quad B = \frac{\partial^2 f}{\partial x \partial y}(1,1) = -\frac{1}{2}, \quad C = \frac{\partial^2 f}{\partial y^2}(1,1) = -\frac{5}{2}$$

de donde $B^2 - AC = \frac{13}{2} > 0$ y entonces, la función f tiene un punto de ensilladura en el punto (1, 1), el cual vale $f(1,1) = \frac{1}{2}\ln 2 + \frac{5\pi}{4} - 1$. ■

Ejemplo 4. Consideremos la función $z = f(x, y)$ dada implícitamente en la expresión

$$F(x,y,z) = x^3 + y^3 + z^3 - 3x - 3y + z + 4 = 0$$

Queremos estudiar sus extremos locales. Los puntos críticos se obtienen al resolver el sistema $\frac{\partial z}{\partial x} = 0$, $\frac{\partial z}{\partial y} = 0$. Estas derivadas las obtenemos con las fórmulas establecidas en el teorema 3.4.2.

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{3x^2-3}{3z^2+1} = 0$$

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{3y^2-3}{3z^2+1} = 0$$

Vemos así que existen cuatro puntos críticos, $\mathbf{p}_1 = (1,1)$, $\mathbf{p}_2 = (1,-1)$, $\mathbf{p}_3 = (-1,1)$, $\mathbf{p}_4 = (-1,-1)$.

Obtengamos las derivadas de segundo orden de la función $z = f(x, y)$.

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(-\frac{3x^2-3}{3z^2+1}\right) = -\frac{(3z^2+1)(6x) - (3x^2-3)\left(6z\frac{\partial f}{\partial x}\right)}{(3z^2+1)^2}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}\left(-\frac{3x^2-3}{3z^2+1}\right) = \frac{3x^2-3}{(3z^2+1)^2}6z\frac{\partial z}{\partial y}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(-\frac{3y^2-3}{3z^2+1}\right) = -\frac{(3z^2+1)(6y) - (3y^2-3)\left(6z\frac{\partial f}{\partial y}\right)}{(3z^2+1)^2}$$

Investiguemos el punto $\mathbf{p}_1 = (1, 1)$. En este caso, el valor de $z = f(1, 1)$ se obtiene de $F(1, 1, z) = 0$, o sea de $z^3 + z = 0$. La única raíz (real) de esta ecuación es $z = 0$. Entonces

$$A = \frac{\partial^2 f}{\partial x^2}(1, 1, 0) = -6, \quad B = \frac{\partial^2 f}{\partial x \partial y}(1, 1, 0) = 0, \quad C = \frac{\partial^2 f}{\partial y^2}(1, 1, 0) = -6$$

Se tiene que $B^2 - AC = -36 < 0$ y $A = -6 < 0$, por lo que concluimos que la función $z = f(x, y)$ tiene un máximo local en $\mathbf{p}_1 = (1, 1)$ que vale $z = 0$.

Para el punto $\mathbf{p}_2 = (1, -1)$, la z correspondiente se obtiene de $F(1, -1, z) = z^3 + z + 4 = 0$. Sea $\bar{z}$ la (única) raíz real de esta ecuación ($\bar{z} \approx -1.378797$). Tenemos

$$A = \frac{\partial^2 f}{\partial x^2}(1, -1, \bar{z}) = -\frac{6}{3\bar{z}^2 + 1}, \quad B = \frac{\partial^2 f}{\partial x \partial y}(1, -1, \bar{z}) = 0$$
$$C = \frac{\partial^2 f}{\partial y^2}(1, -1, \bar{z}) = -\frac{6}{3\bar{z}^2 + 1}$$

se tiene que $B^2 - AC = \frac{36}{(3\bar{z}^2+1)^2} > 0$, por lo que la función $z = f(x, y)$ tiene un punto de ensilladura en el punto $\mathbf{p}_2 = (1, -1)$ que vale $\bar{z} \approx -1.378797$.

Para el punto $\mathbf{p}_3 = (-1, 1)$, la z correspondiente se obtiene de $F(-1, 1, z) = z^3 + z + 4 = 0$. Este es el mismo valor de $\bar{z} = f(-1, 1) \approx -1.378797$. Tenemos

$$A = \frac{\partial^2 f}{\partial x^2}(-1, 1, \bar{z}) = \frac{6}{3\bar{z}^2 + 1}, \quad B = \frac{\partial^2 f}{\partial x \partial y}(-1, 1, \bar{z}) = 0$$
$$C = \frac{\partial^2 f}{\partial y^2}(-1, 1, \bar{z}) = -\frac{6}{3\bar{z}^2 + 1}$$

Se tiene que $B^2 - AC = \frac{36}{(3\bar{z}^2+1)^2} > 0$, por lo que también en el punto $\mathbf{p}_3 = (-1, 1)$ la función $z = f(x, y)$ tiene un punto de ensilladura, que vale $\bar{z} = -1.378797$.

Por último, para el punto $\mathbf{p}_4 = (-1, -1)$, la z correspondiente la obtenemos de $F(-1, -1, \bar{z}) = z^3 + z + 8 = 0$. Sea z^* la (única) raíz real de esta ecuación ($z^* \approx -1.83375$). Tenemos

$$A = \frac{\partial^2 f}{\partial x^2}(-1, -1, z^*) = \frac{6}{3z^* + 1}, \quad B = \frac{\partial^2 f}{\partial x \partial y}(-1, -1, z^*) = 0$$
$$C = \frac{\partial^2 f}{\partial y^2} = \frac{6}{3z^* + 1}$$

Se tiene $B^2 - AC = -\frac{36}{(3z^*+1)^2} < 0$ y $A = \frac{6}{3z^*+1} > 0$, por lo que la función $z = f(x, y)$ tiene un mínimo local en el punto $\mathbf{p}_4 = (-1, -1)$ que vale $z^* = f(-1, -1) \approx -1.83375$. ■

Ejemplo 5. Se quiere construir una caja rectangular abierta cuyo volumen sea de 100 cm³. Nos preguntamos por las dimensiones de la caja, con cuyo volumen tenga la menor superficie posible (figura 3).

Sean x, y, z las dimensiones de la caja, como se muestra en la figura. Su superficie lateral es $2xz + 2yz + xy$. Se quiere entonces determinar el mínimo de la superficie $g(x, y, z) = 2xz + 2yz + xy$ cuando x, y, z satisfacen la relación $xyz = 100$. Esta relación la incorporamos a la función g y obtenemos, por ejemplo, la función

$$f(x, y) = \frac{200}{x} + \frac{200}{y} + xy$$

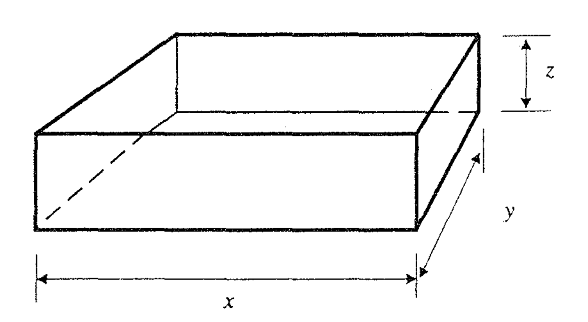


Figura 3. La caja del ejemplo 5.

que describe exactamente la situación de nuestro problema: nos da el valor de la superficie lateral de la caja de capacidad 100 cm^3 en términos de las dimensiones de la base (la altura queda determinada por $z = \frac{100}{xy}$). Obtengamos el punto crítico de f.

$$\frac{\partial f}{\partial x} = -\frac{200}{x^2} + y = 0, \qquad \frac{\partial f}{\partial y} = -\frac{200}{y^2} + x = 0$$

De aquí se obtiene el único punto crítico $\left(\sqrt[3]{200}, \sqrt[3]{200}\right)$. El mismo contexto del problema nos dice que este punto crítico debe ser el mínimo que estamos procurando. Sin embargo, podemos verificarlo con el criterio estudiado en esta sección. Tenemos

$$\frac{\partial^2 f}{\partial x^2} = \frac{400}{x^3}, \qquad \frac{\partial^2 f}{\partial x \partial y} = 1, \qquad \frac{\partial^2 f}{\partial y^2} = \frac{400}{y^3}$$

de donde, en el punto $\left(\sqrt[3]{200}, \sqrt[3]{200}\right)$ se obtienen los valores $A = 2$, $B = 1$, $C = 2$. Así, pues, $B^2 - AC = -3 < 0$, y $A = 2 > 0$, por lo que, en efecto, se tiene un mínimo para la función f en $x = \sqrt[3]{200}$, $y = \sqrt[3]{200}$. La altura de la caja debe ser $z = \frac{100}{(\sqrt[3]{200})^2} = \sqrt[3]{25}$. Concluimos entonces que la caja de dimensiones $x = y = \sqrt[3]{200}$ con $z = \sqrt[3]{25}$, tiene por volumen 100 cm^3 y por superficie, la mínima posible $f\left(\sqrt[3]{200}, \sqrt[3]{200}, \sqrt[3]{25}\right) = 3(200)^{2/3}$ cm^2. ∎

Ejemplo 6. Consideremos el plano $x + y + z = 3$ y el punto $\mathbf{p} = (2, 3, 1)$ (que no se encuentra sobre el plano). Queremos calcular la distancia más corta que existe entre el punto $\mathbf{p}$ y los puntos del plano. A esta distancia se le llama "distancia del punto $\mathbf{p}$ al plano". Sabemos que la distancia de $\mathbf{p}$ al punto (x, y, z) viene dada por

$$d = \sqrt{(x-2)^2 + (y-3)^2 + (z-1)^2}$$

Al incorporar la condición de que el punto (x, y, z) debe ser un punto del plano $x + y + z = 3$, nos queda la función

$$\begin{aligned} f(x, y) &= \sqrt{(x-2)^2 + (y-3)^2 + (3 - x - y - 1)^2} \\ &= \sqrt{(x-2)^2 + (y-3)^2 + (2 - x - y)^2} \end{aligned}$$

Esta es la función que necesitamos estudiar en nuestro problema; ella nos da el valor de la distancia del punto (x, y, z) en el plano $x + y + z = 3$ al punto $\mathbf{p} = (2, 3, 1)$. Queremos obtener el mínimo de esta función.

Los puntos críticos los obtenemos de

$$\frac{\partial f}{\partial x} = \frac{2(x-2) - 2(2-x-y)}{2\sqrt{(x-2)^2 + (y-3)^2 + (2-x-y)^2}} = 0$$
$$\frac{\partial f}{\partial y} = \frac{2(y-3) - 2(2-x-y)}{2\sqrt{(x-2)^2 + (y-3)^2 + (2-x-y)^2}} = 0$$

Es decir, del sistema $2x + y = 4$, $x + 2y = 5$, el cual tiene la única solución $x = 1$, $y = 2$. Por las condiciones del problema, es claro que este punto crítico es un mínimo para la función f, el cual además de ser un mínimo *local*, es de hecho un mínimo *global* para la función (es decir, en este punto la función f toma el valor menor que puede tomar en *todos* los puntos de su dominio). Sin embargo, podemos convencernos de esta situación calculando las derivadas parciales de segundo orden de f en $(1, 2)$; se obtiene que $A = \frac{2}{\sqrt{3}}$, $B = \frac{1}{\sqrt{3}}$, $C = \frac{2}{\sqrt{3}}$, de donde $B^2 - AC = -1 < 0$ y $A = \frac{2}{\sqrt{3}} > 0$, lo que confirma nuestra afirmación. Tenemos pues que el punto del plano $x + y + z = 3$ que está más cerca de $\mathbf{p}$ corresponde a $x = 1$, $y = 2$, $z = 3 - 1 - 2 = 0$. La distancia más corta es entonces,

$$d = f(1, 2) = \sqrt{3}$$

hecho que se confirma con la conocida fórmula de geometría analítica que establece que la distancia entre el punto $\mathbf{p} = (x_0, y_0, z_0)$ y el plano $Ax + By + Cz = D$ la dada

$$d = \frac{|Ax_0 + By_0 + Cz_0 - D|}{\sqrt{A^2 + B^2 + C^2}}$$

fórmula que se puede obtener siguiendo el mismo procedimiento que presentamos en este ejemplo (ver también ejemplo 4 de la próxima sección). ■

Apéndice El método de mínimos cuadrados

En este apéndice expondremos el importante método de mínimos cuadrados, que se aplica para ajustar rectas (en general, la idea del método permite ajustar también otro tipo de curvas) a una serie de datos presentados como puntos del plano.

Supongamos que se tienen los siguientes datos para las variables x y y

x_1	x_2	$\cdots$	x_n
y_1	y_2		y_n

Esta situación se puede presentar en estudios experimentales, donde se estudia la variación de cierta magnitud x (por ejemplo, temperatura, etc.) en función de otra magnitud y. De manera teórica, se espera que la relación entre estas variables sea lineal, es decir, que haya una expresión del tipo $y = mx + b$ que las relacione. El problema es determinar los valores de m y b que determinaron esta funcionalidad lineal entre x y y.

Si las mediciones hechas para x y y fueron exactas, todos los datos (x_i, y_i), $i = 1, 2, \ldots, n$ obtenidos representarían puntos del plano colocados (todos ellos) sobre una misma línea recta y, de hecho, bastarían dos de ellos para determinar la pendiente m y la ordenada al origen b de *la* recta

que determina la relación funcional entre estas variables. Sin embargo, en la práctica, resulta muy poco probable que esto ocurra. Ciertamente, al colocar todos los puntos (x_i, y_i), $i = 1, 2, \ldots, n$ en el plano veríamos que están distribuidos "en los alrededores de alguna recta", la cual se podría trazar sencillamente con una regla, "al tanteo", tratando de que todos quedaran lo "más cercanos posible" a la recta.

El método de mínimos cuadrados nos proporciona un criterio con el cual podremos obtener *la mejor* recta que representa a los puntos dados. Digamos que $y = mx + b$ es esa recta. Como dijimos, sería deseable que se tuviera $y_i = mx_i + b$ para todos los puntos (x_i, y_i) de $i = 1, 2, \ldots, n$. Sin embargo, como, en general, $y_i \neq mx_i + b$, se pide que la suma de los cuadrados de las diferencias (las desviaciones) $y_i - (mx_i + b)$ sea la menor posible. Es decir, se pide que

$$S = (y_1 - (mx_1 + b))^2 + (y_2 - (mx_2 + b))^2 + \cdots + (y_n - (mx_n + b))^2 =$$
$$= \sum_{i=1}^{n} (y_i - (mx_i + b))^2$$

sea lo más pequeña posible. Los valores de m y b que cumplan con esta propiedad, determinan la recta $y = mx + b$ que mejor representa el comportamiento lineal de los puntos (x_i, y_i)

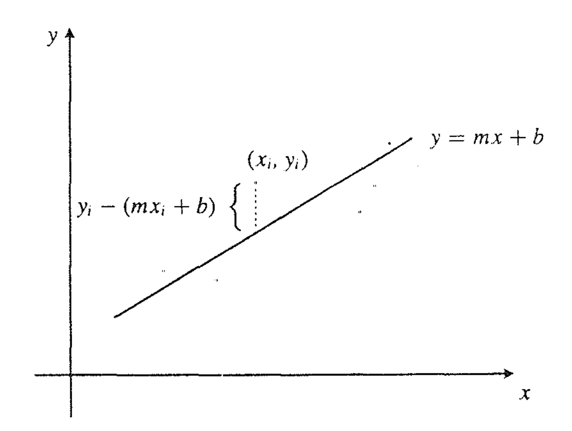


Figura 4. Una recta ajustada.

Consideremos entonces la función f de las variables m y b dada por

$$f(m, b) = \sum_{i=1}^{n} (y_i - (mx_i + b))^2$$

los puntos críticos de esta función se obtienen al resolver el sistema

$$\frac{\partial f}{\partial m} = \sum_{i=1}^{n} 2(y_i - mx_i - b)(-x_i) = -2\sum_{i=1}^{n} x_i(y_i - mx_i - b) = 0$$
$$\frac{\partial f}{\partial b} = \sum_{i=1}^{n} 2(y_i - mx_i - b)(-1) = -2\sum_{i=1}^{n} (y_i - mx_i - b) = 0$$

De la segunda ecuación obtenemos

$$\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} b = 0$$

de donde

$$b = \frac{1}{n}\sum_{i=1}^{n} y_i - m\left(\frac{1}{n}\sum_{i=1}^{n} x_1\right)$$

Llamemos $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. Estas son las medias aritméticas de los valores x_i y y_i, respectivamente. Entonces

$$b = \bar{y} - m\bar{x}$$

Sustituyendo en la ecuación $\frac{\partial f}{\partial m} = 0$, nos queda que

$$\sum_{i=1}^{n} x_i(y_i - mx_i - (\bar{y} - m\bar{x})) = 0$$

de donde se obtiene

$$m = \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sum_{i=1}^{n} x_i(x_i - \bar{x})}$$

En resumen, la función $f(m, b) = \sum_{i=1}^{n}(y_i - mx_i - b)^2$ tiene un (único) punto crítico para

$$m = \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sum_{i=1}^{n} x_i(x_i - \bar{x})}, \quad b = \bar{y} - m\bar{x}$$

Debe ser claro, por el contexto del problema, que en este punto crítico se tiene un mínimo local para la función f, el cual, de hecho, es un mínimo global de f. Sin embargo, para liquidar cualquier brizna de escepticismo en el lector, comprobaremos que, efectivamente, se tiene un mínimo, usando el criterio con las segundas derivadas parciales de la función f. Se tiene

$$A = \frac{\partial^2 f}{\partial m^2} = -2\sum_{i=1}^{n} -x_i^2 = 2\sum_{i=1}^{n} x_i^2$$

$$B = \frac{\partial^2 f}{\partial m \partial b} = -2\sum_{i=1}^{n} -x_i = 2\sum_{i=1}^{n} x_i$$

$$C = \frac{\partial^2 f}{\partial b^2} = -2\sum_{i=1}^{n}(-1) = 2n$$

Se ve que $A > 0$. Resta verificar que $B^2 - AC < 0$, o sea que

$$\left(2\sum_{i=1}^{n} x_i\right)^2 - \left(2\sum_{i=1}^{n} x_i^2\right)(2n) < 0$$

Esta desigualdad es equivalente a

$$\left(\sum_{i=1}^{n} x_i\right)^2 < n \sum_{i=1}^{n} x_i^2$$

la cual no es más que la desigualdad de Cauchy-Schwarz aplicada a los vectores $(1, 1, \ldots, 1)$ y $(x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$ $\big[$Recuerde: si $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ son dos vectores en $\mathbb{R}^n$, entonces se tiene que $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$, donde $\mathbf{x} \cdot \mathbf{y}$ representa el producto punto de los vectores $\mathbf{x}$ y $\mathbf{y}$, y $\|\mathbf{x}\| = \sqrt{\sum x^2}$. Esta desigualdad es una igualdad si y sólo si los vectores x y y son linealmente dependientes (ver ejercicio 2 de la sección 2, capítulo 1)$\big]$. En nuestro caso, poniendo $\mathbf{y} = (1, 1, \ldots, 1)$, que es linealmente independiente de $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ (¿por qué?) obtenemos $|\sum x_i| \leq \sqrt{\sum 1}\sqrt{\sum x_i^2}$, o sea $\left(\sum x_i\right)^2 < n \sum x_i^2$, como se quería). Así pues, $B^2 - AC < 0$ y $A > 0$, lo cual nos dice que la función f posee un mínimo local en el punto crítico estudiado.

En resumen, la *mejor* recta (en el sentido de los mínimos cuadrados) que describe el comportamiento lineal de los puntos (x_i, y_i), $i = 1, 2, \ldots, n$ es la recta $y = mx + b$, donde

$$m = \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sum_{i=1}^{n} x_i(x_i - \bar{x})}, \qquad b = \bar{y} - m\bar{x}$$

Veamos un ejemplo.

Se obtuvieron experimentalmente los siguientes valores de las variables x y y, los cuales se sabe que guardan entre sí una relación lineal.

x	1.0	2.0	3.0	4.0
y	1.4	1.1	0.7	0.1

Veamos cuál es la mejor recta que ajusta estos datos, según el método de mínimos cuadrados.

Tenemos

$$\bar{x} = \frac{1 + 2 + 3 + 4}{4} = 2.5$$
$$\bar{y} = \frac{1.4 + 1.1 + 0.7 + 0.1}{4} = 0.825$$

Aplicando la fórmula obtenida para m y b se obtiene

$$\begin{aligned} m &= \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sum_{i=1}^{n} x_i(x_i - \bar{y})} \\ &= \frac{1(1.4 - 0.825) + 2(1.1 - 0.825) + 3(0.7 - 0.825) + 4(0.1 - 0.825)}{1(1 - 2.5) + 2(2 - 2.5) + 3(3 - 2.5) + 4(4 - 2.5)} \\ &= \frac{-2.15}{5} = -0.43 \\ b &= \bar{y} - m\bar{x} = 0.825 - (-0.43)(2.5) = 1.9 \end{aligned}$$

Entonces, la recta $y = -0.43x + 1.9$ es la que mejor ajusta a los datos proporcionados. Véase la tabla siguiente

x	y	y calculada con $y=-0.43x+1.9$	Desviación
1	1.4	1.47	−0.07
2	1.1	1.04	0.06
3	0.7	0.61	0.09
4	0.1	0.18	−0.08

La suma de las diferencias de la y real con la y predicha por la ecuación obtenida es $-0.07+0.06+0.09-0.08=0$. Es decir, nuestra recta efectivamente *compensa* los puntos que quedaron por encima de ella con los puntos que quedaron por debajo.

Gráficamente esto se ve como

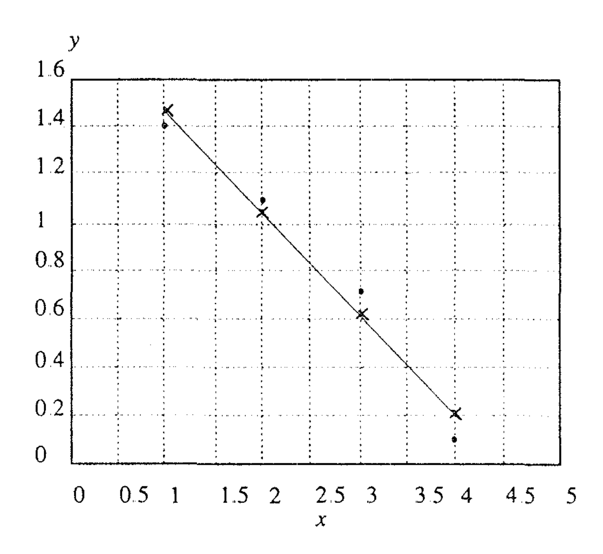


Figura 5. La mejor recta que ajusta los datos del ejemplo. ■

Ejercicios (Capítulo 4, Sección 4)

En los ejercicios 1–27 determine la naturaleza de los puntos críticos de las funciones dadas (si los hay).

1. $f(x, y) = 2x^2 + 3y^2 + xy - 5x + 6y - 27$

2. $f(x, y) = xy + x + y + 1$

3. $f(x, y) = x^2y^2 + 2x + 2y + 1$

4. $f(x, y) = x^3 + y^3 - 3x - 3y - 2$

5. $f(x, y) = x^4 + 3y^3 - 2x^2 - 3y - 1$

6. $f(x, y) = xy(1 - x - y)$

7. $f(x, y) = x\exp(-x^2 - y^2)$

8. $f(x, y) = y \exp(-x^2 - y^2)$

9. $f(x, y) = x \ln y$

10. $f(x, y) = x \ln y - x$

11. $f(x, y) = 3y \ln x - 4x + 2$

12. $f(x, y) = (y^2 - 1) \ln x - x$

13. $f(x, y) = \ln(1 + x^2 + y^2)$

14. $f(x, y) = y \arctan x - 2x - y + 1$

15. $f(x, y) = \cosh x + \cosh y$

16. $f(x, y) = \cos x + \cos y$

17. $f(x, y) = \operatorname{sen} x \cos y$

18. $f(x, y) = x + \dfrac{y}{x} + y$

19. $f(x, y) = x + \dfrac{y}{x} - 0.5y^2$

20. $f(x, y) = (x^2 - 1) \exp(-y^2)$

21. $f(x, y) = (x^2 - 1) \exp(-y) - y$

22. $f(x, y) = \arctan(x^2) + \arctan(y^2)$

23. $f(x, y) = \dfrac{1}{\cosh x} + 1 - y^2$

24. $f(x, y) = \operatorname{sech} x + \operatorname{sech} y$

25. $f(x, y) = \ln\left(1 + \sqrt{1 + x^2 + y^2}\right)$

26. $f(x, y) = \ln(1 + \exp(-x^2 - y^2))$

(*) 27. $f(x, y) = \left(\ln(1 + \exp(-x^2) + \exp(-y^2))\right)^{-1}$

En cada uno de los ejercicios 28–36 se dan funciones $z = f(x, y)$ en forma implícita. Determine la naturaleza de sus puntos críticos.

28. $-x^2 + 2y^2 + 3z^2 + 4xy + 5xz + 6yz + 7x + 8y + 9z + 115/2 = 0$

29. $-2x^2 + 2z^2 + 3xy + xz + yz - 13/9 = 0$

30. $3x^2 + 2y^2 + z^2 + 3xy + 2xz - yz + x + y + 2 = 0$

31. $2x^2 + 2y^2 + z^2 + xz + 2yz + x + 2y + 3z = 0$

32. $11x^2 + 9y^2 + 7z^2 + 5xy + 3xz + yz - 530 = 0$

33. $3x^2 + 4y^2 + 5z^2 + xz + 2yz + 3x + 4y + 5z + 7/4 = 0$

34. $-4x^2 - y^2 - 3z^2 + 3yz + 3y + 6z - 36 = 0$

35. $-x^2 + 2z^2 - xy + xz + x + 2y + z = 0$

36. $2x^2 - y^2 - z^2 - 3xy + 2xz + yz + x + y + z + 220 = 0$

37. Sean $g, h: \mathbb{R} \to \mathbb{R}$ dos funciones diferenciables dos veces. Suponga que g tiene solamente un extremo local en x_0, el cual es un mínimo que vale $g(x_0) = a$, y que h tiene también solamente un extremo local en x_1, el cual es un mínimo que vale $h(x_1) = b$. Suponga además que las segundas derivadas de g y h son no nulas en sus puntos críticos, y que las gráficas de estas funciones no cruzan al eje x. Determine la naturaleza de los puntos críticos de las siguientes funciones $f: \mathbb{R}^2 \to \mathbb{R}$.

a. $f(x, y) = g(x) + h(y)$

b. $f(x, y) = g(x)h(y)$, en los siguientes casos:

b1. $a > 0, b > 0$, **b2.** $a > 0, b < 0$

b1. $a < 0, b > 0$, **b4.** $a < 0, b < 0$

c. $f(x, y) = (g(x))^2 + (h(y))^2$, en los siguientes casos:

c1. g y h son funciones positivas.

c2. g es positiva y h es negativa.

c3. g es negativa y h es positiva.

c4. g y h son negativas.

38. Determine la naturaleza de los puntos críticos de la función $f(x, y) = (x^2 + 1)^2 + (y^2 - y + 1)^2$. (Concilie con el resultado del inciso $c1$) del ejercicio anterior).

39. Determine la naturaleza de los puntos críticos de la función $f(x, y) = (x^2 + 1)^2 + (-e^{-y^2} - 3)^2$. (Concilie con el resultado del inciso $c2$) del ejercicio anterior).

Los ejercicios 40–46 se refieren al método de mínimos cuadrados estudiado en el apéndice de esta sección.

En cada uno de los ejercicios 40–42, obtenga la mejor recta, según el criterio de mínimos cuadrados, que ajusta los datos ahí presentados. En cada caso, compare el valor de la variable y, correspondiente a cada valor de x dado, con el obtenido por la recta calculada.

40.

x	1	2	2.5	3.01	4	5.7
y	0.9	2.11	2.66	2.92	4	5.5

41.

x	0.21	0.67	1.89	3.33	4.44	8.78
y	−0.4	0.91	4.50	9.11	12.1	26.5

42.

x	−2.1	0	2.51	3.33	6.71	7.09
y	−1.1	3	7.99	9.62	16.5	25.1

43. En determinadas condiciones, se puede considerar que la reacción de oxidación de tiourea por hexacano ferrato (III) en medio alcalino

$$NH_2CSNH_2 + 10(OH)^- + 8(Fe(CN)_6)^{3-} \longrightarrow NH_2CONH_2 + SO_4^{2-} + 8(Fe(CN)_6)^{4-} + 5H_2O$$

es una reacción de primer orden, para la cual se cumple que

$$C = C_0 e^{-kt}$$

donde C es la concentración molar de hexaciano a un tiempo t, C_0 es la concentración inicial de este reactivo (concentración al tiempo $t = 0$), y k es la llamada *constante de velocidad*

de reacción, parámetro cuyo conocimiento reporta una gran cantidad de información sobre la misma. Obsérvese que la expresión anterior se puede escribir como

$$\ln C = \ln C_0 - kt$$

de donde queda clara la dependencia lineal entre $\ln C$ y t. El valor de k es una función del pH de la solución reaccionante, de la fuerza iónica y de la temperatura a la que se efectúa la reacción. A continuación se muestran varias tablas de datos obtenidos experimentalmente a diferentes condiciones durante la reacción, en las que se dan los valores de C (en moles por litro) a diferentes tiempos. En cada caso, calcule el valor de k.

a. pH = 9.0, fuerza iónica = normal, temperatura = 40 °C.

t (mín)	3	15	40	80	100	120
C ($\times 10^4$) (mol/l)	7.847	7.406	5.721	3.486	2.867	2.704

b. pH = 9.0, fuerza iónica = normal, temperatura = 50°C.

t (mín)	3	15	40	80	100	120
C ($\times 10^4$) (mol/l)	7.406	6.055	2.956	1.800	1.586	1.485

c. pH = 11.0, fuerza iónica = normal, temperatura = 40 °C.

t (mín)	2	7	9	12	16	19
C ($\times 10^4$) (mol/l)	6.217	4.802	4.032	3.336	2.485	1.791

d. pH = 11.0, fuerza iónica = normal, temperatura = 50 °C.

t (mín)	1	2	3	4	5.5	8
C ($\times 10^4$) (mol/l)	5.911	4.550	3.033	2.152	1.676	1.356

44. (Regresión lineal en tres variables). Supongamos que se tienen n datos de 3 variables x, y, z, a saber (x_i, y_i, z_i), $i = 1, 2, \ldots, n$, de las cuales se sabe que deben guardar una relación lineal del tipo $z = Ax + By + C$. Se quieren determinar entonces los coeficientes A, B y C que hacen que los valores de z calculados como $z = Ax + By + C$, ajusten lo mejor posible (en el sentido de los mínimos cuadrados) los datos proporcionados. Desde el punto de vista geométrico, se trata de determinar el plano $z = Ax + By + C$ que mejor ajuste los n puntos dados $(x_i, y_i, z_i) \in \mathbb{R}^3$, $i = 1, 2, \ldots, n$. Tomando entonces la suma de los cuadrados de las diferencias de z_i con $Ax_i + By_i + C$ = valor de z en (x_i, y_i) según el plano que se quiere determinar, se trata de hallar los valores de A, B y C que hagan que esta suma sea la menor posible. Es decir, se trata de hallar los valores de A, B y C que hagan que

$$S = S(A, B, C) = \sum_{i=1}^{n} (z_i - (Ax_i + By_i + C))^2$$

sea mínima. Demuestre que, resolviendo el sistema $\frac{\partial S}{\partial A} = 0$, $\frac{\partial S}{\partial B} = 0$, $\frac{\partial S}{\partial C} = 0$, se obtiene, como único punto crítico, A, B, C, donde

$$C = \bar{z} - A\bar{x} - B\bar{y}$$

y A y B son las soluciones del sistema

$$\begin{bmatrix} \sum x_i^2 - n\bar{x}^{-2} & \sum x_i y_i - n\bar{x}\bar{y} \\ \sum x_i y_i - n\bar{x}\bar{y} & \sum y_i^2 - n\bar{y}^2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum x_i(z_i - \bar{z}) \\ \sum y_i(z_i - \bar{z}) \end{bmatrix}$$

($\bar{x}$, $\bar{y}$, $\bar{z}$ son las medias aritméticas de los valores dados x_i, y_i, z_i, $i = 1, 2, \ldots, n$, las sumas son de $i = 1$ a $i = n$). Es claro que este punto crítico debe tratarse de un mínimo para la función $S(A, B, C)$ (¿por qué?). Tome este resultado como base para obtener el mejor ajuste lineal del tipo $z = Ax + By + C$ para cada uno de los siguientes conjuntos de datos.

a.

x	0	0.8	1	3
y	0.1	−0.5	−1	−2
z	1.1	1.3	1	2

b.

x	1.8	4.2	−0.3	2	3
y	0.5	2.1	3.2	−0.4	1
z	2	1.8	−2.7	3.5	2.5

c.

x	−0.1	2	1.4	−3.5	0	−1
y	3.2	0	5.3	−0.4	−1.6	2
z	−0.2	2.5	−0.42	0	2.35	0.1

d.

x	0.4	2.3	−0.2	1.3	0.
y	−1	1.7	0.8	1.1	3
z	−4.5	2	1.5	0.9	8.05

e.

x	0	−1	2	3
y	5	0.4	−1.5	0.5
z	0.9	0	−4.7	−5.1

45. (Ajuste cuadrático). Supongamos un conjunto de datos de las variables x, y, a saber (x_i, y_i), $i = 1, 2, \ldots, n$, las cuales se sabe que guardan una relación cuadrática del tipo $y = ax^2 + bx + c$. Se trata de determinar los coeficientes a, b, c que mejor ajustan (según el criterio de los mínimos cuadrados) tal relación cuadrática. Desde el punto de vista geométrico, se trata de determinar la parábola $y = ax^2 + bx + c$, que mejor ajusta al conjunto de los n puntos del plano (x_i, y_i). Demuestre que este problema se puede reducir al considerado en el ejercicio anterior, calculando la función lineal $z = Ax + By + C$, que mejor ajusta el comportamiento lineal de las variables x, y $(= x^2)$, z = la ordenada de la parábola por determinar. Tome esta idea para obtener la parábola que mejor ajusta cada uno de los siguientes conjuntos de datos.

a.

x	1	−1	2	3
y	3.9	6	9.1	18

b.

x	0	1	−1	2	−2
y	−1.1	6.1	−11.95	8.95	−27

c.

x	0	1.5	−1	2	−0.5
y	4.1	1.7	7.9	2.2	5.9

46. (Regresión lineal en $N + 1$ variables). Supongamos un conjunto de m datos de las n variables $x_1, x_2, \ldots, x_n$, y, digamos $(x_{1_i}, x_{2_i}, \ldots, x_{n_i}, y_i)$ $i = 1, 2, \ldots, m$, las cuales se sabe que deben guardar una relación del tipo $y = A_1x_1 + A_2x_2 + \cdots + A_nx_n + B$. Se trata de determinar los coeficientes A_i y B que mejor ajustan esta relación (según el criterio de los mínimos cuadrados) con los datos presentados. Tomando entonces la función $S = S(A_1, \ldots, A_n, B)$ = suma de los cuadrados de las diferencias entre los valores y_i dados y los valores $(A_1x_{1_i} + \ldots + A_nx_{n_i} + B)$, es decir

$$S = S(A_1, \ldots, A_n, B) = \sum_{i=1}^{m} \left(y_i - (A_1x_{1_i} + \ldots + A_nx_{n_i} + B)\right)^2$$

se trata de determinar los valores de $A_1, \ldots, A_n$, B que hagan que S sea mínima. Demuestre que el sistema $\frac{\partial S}{\partial A_i} = 0$, $\frac{\partial S}{\partial B} = 0$, $i = 1, 2, \ldots, n$, tiene por solución a

$$B = \bar{y} - A_1\bar{x}_1 - \cdots - A_n\bar{x}_n$$

donde $A_1, \ldots, A_n$ son las soluciones del sistema $M\times = N$, siendo

$$M = \begin{bmatrix} \sum x_{1_i}^2 - m\bar{x}_1^2 & \sum x_{1_i}x_{2_i} - m\bar{x}_1\bar{x}_2 & \cdots & \sum x_{1_i}x_{n_i} - m\bar{x}_1\bar{x}_n \\ \sum x_{1_i}x_{2_i} - m\bar{x}_1\bar{x}_2 & \sum x_{2_i}^2 - m\bar{x}_2^2 & \cdots & \sum x_{2_i}x_{n_i} - m\bar{x}_2\bar{x}_n \\ & \cdots & & \\ \sum x_{n_i}x_{1_i} - m\bar{x}_1\bar{x}_n & \sum x_{n_i}x_{2_i} - m\bar{x}_2\bar{x}_n & \cdots & \sum x_{n_i}^2 - m\bar{x}_n^2 \end{bmatrix}$$

$$X = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}, \quad N = \begin{bmatrix} \sum x_{1_i}y_i - m\bar{x}_1 y \\ \sum x_{2_i}y_i - m\bar{x}_2 y \\ \vdots \\ \sum x_{n_i}y_i - m\bar{x}_n y \end{bmatrix}$$

(las sumas son de $i = 1$ a $i = m$, $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{y}$, que son las medias aritméticas de los valores correspondientes de estas variables). Siga estas ideas para obtener la mejor aproximación lineal del tipo $y = A_1x_1 + \cdots + A_nx_n + B$, para cada uno de los siguientes conjuntos de datos.

a.

x_1	2	3.1	4.2	6.7	7.5
x_2	1	−0.5	−2	−3.4	−5.1
x_3	0	2.6	3.2	5.1	6.7
y	4	1.7	1.4	4.2	1.6

b.

x_1	1	1.3	2.5	4.8	5.6
x_2	2	2.4	3.7	3.9	6.6
x_3	3	3.7	5.2	5.9	8.4
x_4	4	4.9	6.7	7.7	10.1
y	−0.9	0	2.5	9	12

4.5 Extremos condicionados

Hay un tipo especial de problemas en relación a encontrar los valores máximo y mínimo de una función de varias variables, muy importante por sus aplicaciones, que tiene que ver con un planteamiento como el siguiente: supongamos que queremos encontrar los extremos de la función $z = f(x, y)$, cuando las variables x, y varían en un conjunto determinado de puntos en el plano como podría ser una curva. Es decir, nos interesa obtener cuál es el valor más grande y más pequeño (localmente) de la función $z = f(x, y)$, y en qué puntos se tienen estos valores cuando (x, y) se mueve sobre una curva presentada previamente, digamos que por una ecuación del tipo $g(x, y) = 0$. La diferencia entre este planteamiento del problema de extremos de la función $z = f(x, y)$ con el expuesto en la sección anterior para extremos locales, es que en estos últimos las variables x, y se movían en *todo* el dominio función f, y se procuraban los puntos $(\bar{x}, \bar{y})$ para los que el valor $f(\bar{x}, \bar{y})$

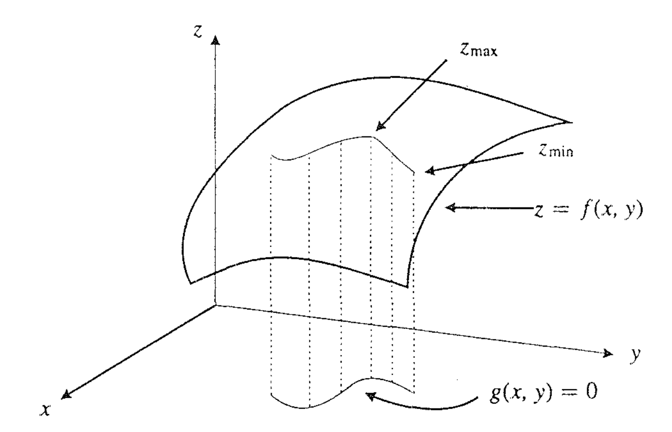


Figura 1. Valor mínimo y máximo de $z = f(x, y)$ con (x, y) variando en la curva C

fuera el mayor o menor de los valores de $f(x, y)$ con (x, y) variando en una bola con centro en $(\bar{x}, \bar{y})$. La novedad ahora es que procuramos los puntos $(\bar{x}, \bar{y})$ para los que el valor $f(\bar{x}, \bar{y})$ es el mayor o menor de los valores de $f(x, y)$, con (x, y) variando en una curva dada $g(x, y) = 0$. Gráficamente (figura 1)

Tomemos el siguiente ejemplo concreto que nos servirá para entender el planteamiento y la propuesta de solución para este tipo de problemas en general.

Damos la curva en el plano

$$(x-3)^2 + \frac{1}{4}(y-4)^2 = 1$$

o bien $g(x, y) = 4x^2 + y^2 - 24x - 8y + 48 = 0$. Esta es una elipse con centro en el punto $(3, 4)$, semiejes 1 y 2, y eje mayor paralelo al eje y. Queremos encontrar qué punto de esta elipse se encuentra más cercano al origen y qué punto se encuentra más alejado de él. Para un punto (x, y) en el plano, la distancia entre él y el origen es $\sqrt{x^2 + y^2}$. Podemos pensar entonces en la "función distancia del punto (x, y) al $(0, 0)$", dada por $f(x, y) = \sqrt{x^2 + y^2}$. Esta es una función de las variables x, y, de la cual nos interesa obtener el máximo y el mínimo, *cuando x, y se mueven sobre la elipse $g(x, y) = 0$* Gráficamente (figura 2)

Diremos entonces que queremos encontrar el máximo y el mínimo de la función $z = f(x, y) = \sqrt{x^2 + y^2}$ sujeta a la restrición $g(x, y) = 4x^2 + y^2 - 24x - 8y + 48 = 0$. Estos valores extremos de la función $z = f(x, y)$ se llaman, en general, ***extremos condicionados***.

Por supuesto que una manera de resolver nuestro ejemplo es despejar, de $g(x, y) = 0$, la variable y en términos de la variable x, obteniendo así una función del tipo $y = \varphi(x)$, y sustituir ésta en la función $z = f(x, y)$, convirtiendo así el problema inicial al de la determinación de extremos de la función $z = f(x, \varphi(x))$, la cual depende de la única variable x. En general este procedimiento queda supeditado a que efectivamente se pueda hacer el despeje mencionado. En este momento nuestro objetivo es investigar cómo atacar el problema sin esta limitante.

Fijemos nuestra atención en las curvas de nivel de la función f; es decir, las curvas $f(x, y) = C$. Estas son los círculos $\sqrt{x^2 + y^2} = C$, que tienen su centro en el origen y radio C. Sea $(\bar{x}, \bar{y})$ el punto en donde se alcanza el mínimo buscado (el cual obviamente existe) y sea $\bar{C} = \sqrt{\bar{x}^2 + \bar{y}^2}$. Observemos con detenimiento la curva de nivel $\sqrt{x^2 + y^2} = \bar{C}$. Esta curva debe ser tangente a la elipse $g(x, y) = 0$ en el punto $(\bar{x}, \bar{y})$. En cierto sentido la función de la recta tangente común

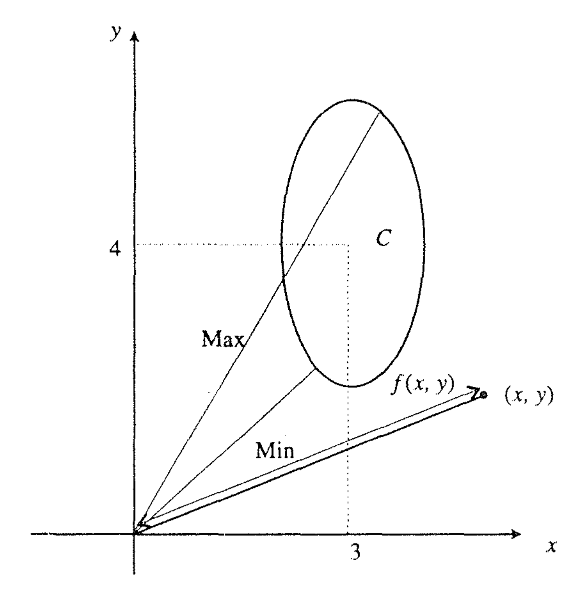


Figura 2. La elipse del ejemplo.

es marcar la separación del punto $(\bar{x}, \bar{y})$ con los demás puntos de la elipse que se encuentran más alejados del origen que $(\bar{x}, \bar{y})$ (que es el más cercano al origen) (figura 3).

Sabemos que el vector grad $f(\bar{x}, \bar{y})$ es un vector perpendicular a la curva de nivel que pasa por $(\bar{x}, \bar{y})$. Entonces éste es un vector perpendicular al círculo $x^2 + y^2 = \bar{C}^2$ en el punto $(\bar{x}, \bar{y})$. Por otra parte, la elipse $g(x, y) = 0$ la podemos ver como una curva de nivel (correspondiente al nivel cero) de la función $z = g(x, y)$. Entonces el vector grad $g(\bar{x}, \bar{y})$ debe ser perpendicular a tal curva en $(\bar{x}, \bar{y})$. En resumen: grad $f(\bar{x}, \bar{y})$ es perpendicular a $f(x, y) = \bar{C}$ en $(\bar{x}, \bar{y})$; grad $g(\bar{x}, \bar{y})$ es perpendicular a $g(x, y) = 0$ en $(\bar{x}, \bar{y})$; las curvas $f(x, y) = \bar{C}$ y $g(x, y) = 0$ son tangentes en $(\bar{x}, \bar{y})$. Conclusión: grad $f(\bar{x}, \bar{y})$ y grad $g(\bar{x}, \bar{y})$ son vectores *colineales*.

Un razonamiento análogo se aplica para el punto (x^*, y^*) en el que se alcanza el máximo.

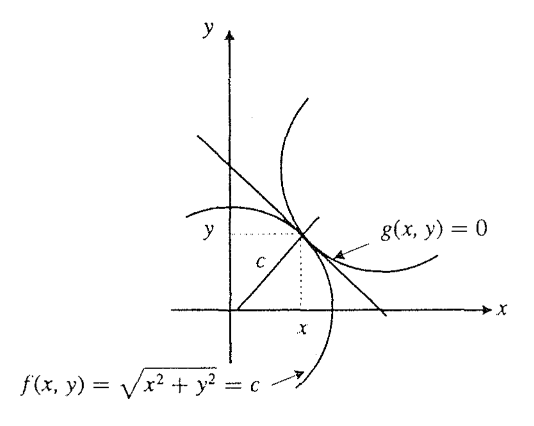


Figura 3. La recta tangente a la elipse y la curva de nivel.

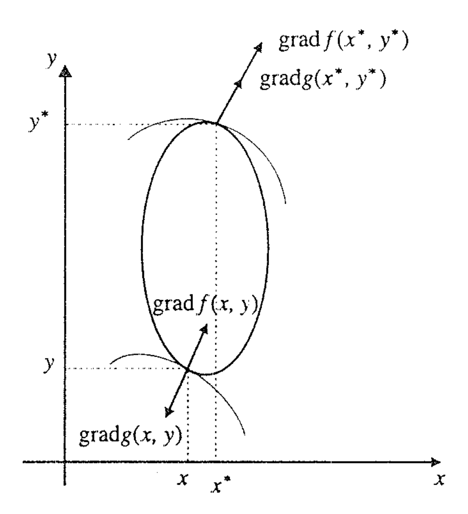


Figura 4. Los vectores grad f y grad g donde se alcanzan los extremos.

Este análisis nos permite hacer la siguiente importante afirmación: en el punto (x, y), donde la función $z = f(x, y)$ alcance su máximo o mínimo condicionado, debe existir una constante λ tal que

$$\text{grad } f(x, y) = \lambda \text{ grad } g(x, y)$$

Esta es precisamente la condición de colinealidad de los vectores grad $f(x, y)$ y grad $g(x, y)$ que hemos concluido ocurre en los puntos del extremo.

Busquemos pues en qué puntos ocurre una relación como la anterior. Se debe tener entonces que

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \lambda\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) \quad \text{o sea} \quad \frac{\partial f}{\partial x} = \lambda\frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda\frac{\partial g}{\partial y}$$

o bien

$$\frac{x}{\sqrt{x^2+y^2}} = \lambda(8x-24), \qquad \frac{y}{\sqrt{x^2+y^2}} = \lambda(2y-8)$$

Tenemos así un sistema de tres ecuaciones (las dos anteriores más la ecuación $g(x, y) = 0$) con tres indeterminadas x, y, λ. Resolviendo este sistema localizamos los puntos (x, y) de la elipse en la que se da la colinealidad de los vectores grad $f(x, y)$ y grad $g(x, y)$. Estos son los puntos donde se encuentran los extremos procurados.

Sin entrar en detalles técnicos (los cuales pueden ser muy complicados, como en este ejemplo que conduce a resolver una ecuación de cuarto grado) sólo diremos que el sistema

$$\begin{aligned} \frac{x}{\sqrt{x^2+y^2}} &= \lambda(8x-24) \\ \frac{y}{\sqrt{x^2+y^2}} &= \lambda(2y-8) \\ 4x^2+y^2-24x-8y+48 &= 0 \end{aligned}$$

tiene solución para $x = 2.49$, $y = 2.198$ y $x = 3.26$, $y = 5.8765$. Viendo que $f(2.49, 2.198) = 3.3213$, $f(3.26, 5.8765) = 6.72018$, debe ser claro que el punto $(2.49, 2.198)$ es el punto de la elipse

más cercano al origen (a una distancia de 3.3213 unidades), en tanto que el punto (3.26, 5.8765) es el que se encuentra más lejos del origen (a una distancia de 6.72018 unidades). Esta es la solución de nuestro problema. El valor de λ que satisface el sistema no interesa mucho; sólo nos dice el factor por el que hay que multiplicar el vector grad $g(x, y)$, para obtener grad $f(x, y)$ en los puntos donde ocurre el extremo. (En nuestro ejemplo tenemos que en (2.49, 2.198) el valor de λ es -0.1837 y en (3.26, 5.8765) su valor es 0.12886, valores cuyo signo coinciden con lo previsto en la figura 4.)

Consideremos ahora, en abstracto, una situación como la mostrada en el ejemplo anterior. Sea la función $z = f(x, y)$ de la cual se quieren obtener sus extremos cuando (x, y) se mueve sobre la curva $g(x, y) = 0$. Supongamos que en el punto $(\bar{x}, \bar{y})$ ocurre uno de los extremos, el máximo. Esto significa que $f(\bar{x}, \bar{y})$ es el valor más grande de los valores de $f(x, y)$ cuando (x, y) se encuentra en una vecindad *sobre la curva* $g(x, y) = 0$ de centro en $(\bar{x}, \bar{y})$ (más aún $f(\bar{x}, \bar{y}) \geq f(x, y)$ cuando (x, y) se encuentra en $B \cap C$ donde B es una bola con centro en $(\bar{x}, \bar{y})$ y C es la curva $g(x, y) = 0$). Consideremos la curva de nivel de $z = f(x, y)$ correspondiente al nivel $C = f(\bar{x}, \bar{y})$. Es decir, consideremos la curva

$$L = \{(x, y) \in \mathbb{R}^2 | f(x, y) = f(\bar{x}, \bar{y})\}$$

La curva L separa (en general) los puntos del plano en los que $f(x, y) > f(\bar{x}, \bar{y})$ de los puntos en los que $f(x, y) < f(\bar{x}, \bar{y})$. ¿Cómo es la curva $g(x, y) = 0$ respecto de la curva L en los alrededores de $(\bar{x}, \bar{y})$ —donde coinciden? Observemos que una situación como la mostrada en la figura 5 no puede ocurrir pues habrá puntos de la curva $g(x, y) = 0$ vecinos a $(\bar{x}, \bar{y})$ para los que $f(x, y) > f(\bar{x}, \bar{y})$.

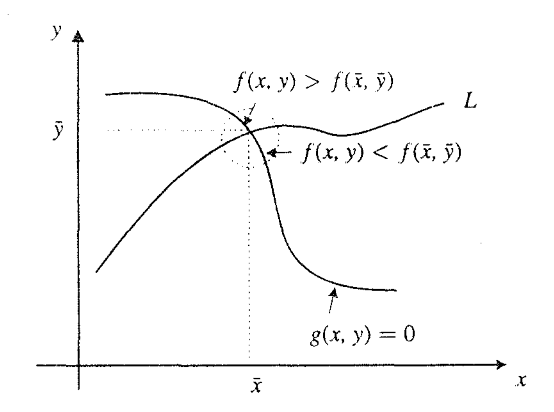


Figura 5. Gráfica de las curvas $g(x, y) = 0$ y L.

Se ve que para que se pueda tener el máximo de $f(x, y)$ en $(\bar{x}, \bar{y})$ toda la curva $g(x, y) = 0$ debe quedar en la región $f(x, y) < f(\bar{x}, \bar{y})$, al menos en una vecindad de $(\bar{x}, \bar{y})$. Sólo así garantizamos que $f(x, y) \leq f(\bar{x}, \bar{y})$ para toda (x, y) en la curva $g(x, y) = 0$ cerca de $(\bar{x}, \bar{y})$. Entonces, en el punto $(\bar{x}, \bar{y})$ debe haber tangencia entre L y $g(x, y) = 0$.

Como L es una curva de nivel de $z = f(x, y)$, se debe tener que el vector grad $f(\bar{x}, \bar{y})$ es perpendicular a L en $(\bar{x}, \bar{y})$. Así mismo, la curva $g(x, y) = 0$ se puede ver como una curva de nivel de la función $z = g(x, y)$, de modo que también grad $g(\bar{x}, \bar{y})$ debe ser perpendicular a $g(x, y) = 0$ en $(\bar{x}, \bar{y})$. Siendo L y $g(x, y) = 0$ tangentes en $(\bar{x}, \bar{y})$, concluimos que los vectores grad $f(\bar{x}, \bar{y})$ y grad $g(\bar{x}, \bar{y})$ son colineales (figura 7)

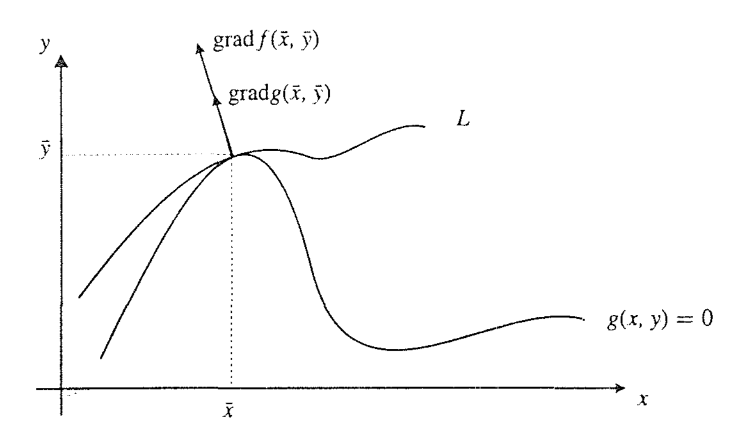


Figura 7. Gráfica que muestra la colinealidad de grad $f(\bar{x}, \bar{y})$ y grad $g(\bar{x}, \bar{y})$.

Es decir, debe ocurrir $\lambda \in \mathbb{R}$ de modo que grad $f(\bar{x}, \bar{y}) = \lambda$ grad $g(\bar{x}, \bar{y})$. En resumen, podemos establecer lo siguiente: una condición necesaria para que la función $z = f(x, y)$ sujeta a la restricción $g(x, y) = 0$ tenga un extremo en el punto (x, y), (donde grad $g(x, y) \neq 0$ (una condición técnica que garantiza que en los alrededores del extremo buscado, la curva $g(x, y) = 0$ se ve como una "curva decente", gráfica de alguna función $y = \varphi(x)$ ó $x = \psi(y)$—lo cual se explica en términos del teorema de la Función Implícita, ver ejemplo 2 de la siguiente sección)) es que exista $\lambda \in \mathbb{R}$ tal que

$$\text{grad } f(x, y) = \lambda \text{ grad } g(x, y)$$

Esta condición, junto con el hecho de que el punto procurado debe ser un punto de la curva $g(x, y) = 0$, nos proporciona el sistema de 3 ecuaciones

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}, \quad g(x, y) = 0$$

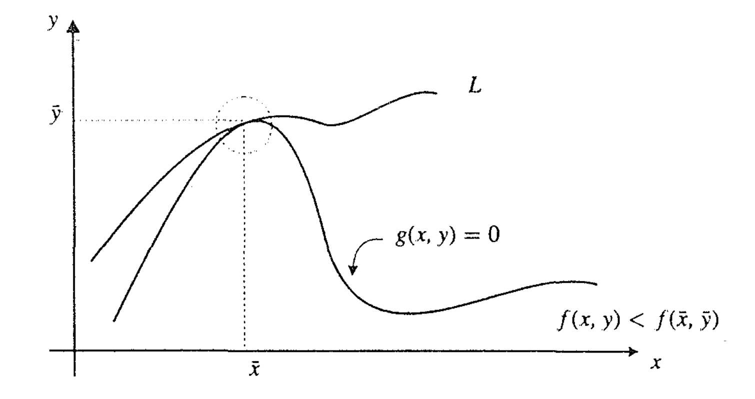


Figura 6. Gráfica de la tangencia de las curvas L y $g(x, y) = 0$.

con tres indeterminadas x, y, λ, el cual, resolviéndolo, nos dará los (posibles, como veremos) extremos condicionados que se buscan.

Una vez localizados los puntos (x, y) que cumplen las condiciones establecidas en el párrafo anterior, muchas veces resulta muy sencillo, ya sea por el contexto físico o geométrico del problema, o bien, por un análisis simple alrededor de esos puntos, establecer si en tales puntos se encuentra efectivamente un máximo o un mínimo. Debemos advertir, sin embargo, que en tales puntos podrá no haber extremo para la función f. En otras palabras, la condición de colinealidad de los vectores grad f y grad g, es necesaria *pero no suficiente* para la existencia de un extremo condicionado para la función f. La siguiente gráfica debe explicar este hecho. En la siguiente sección ahondaremos más a este respecto. Por el momento, nuestros esfuerzos se concentrarán en entender la estrategia de ataque en general, para el problema de extremos condicionados.

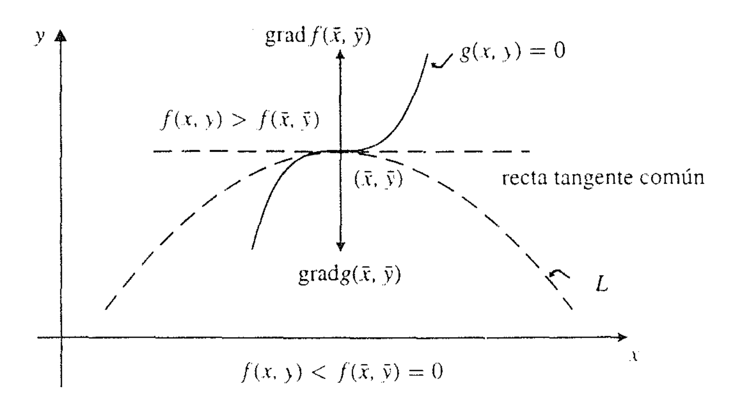


Figura 8. La colinealidad de los vectores grad f y grad g no es suficiente para la existencia de extremos condicionados.

Antes de considerar la situación general, analicemos otro ejemplo concreto.

Supongamos que se quieren obtener los extremos de la función $u = f(x, y, z)$ cuando (x, y, z) varía en una curva C del espacio, descrita como la intersección de dos superficies $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$. Ahora tenemos un problema de extremos condicionados donde se buscan los extremos de una función de tres variables $u = f(x, y, z)$ sujeta a dos restricciones $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$ (figura 9).

Supongamos que en el punto $\mathbf{p} = (\bar{x}, \bar{y}, \bar{z})$ se logra uno de los extremos condicionados procurado, digamos que el máximo. Tomamos una función $\alpha: I \subseteq \mathbb{R} \to \mathbb{R}^3$ cuyas imágenes describen precisamente la curva C (este tipo de funciones son objeto de estudio del próximo capítulo). Es decir

$$\alpha(t) \in C = \{(x, y, z) \in \mathbb{R}^3 | g_1(x, y, z) = 0\} \cap \{(x, y, z) \in \mathbb{R}^3 | g_2(x, y, z) = 0\}$$

para $t \in I$. Digamos que $\bar{t} \in I$ es tal que $\alpha(\bar{t}) = \mathbf{p} = (\bar{x}, \bar{y}, \bar{z})$. Podemos formar la composición $f \circ \alpha: I \subseteq \mathbb{R} \to \mathbb{R}$, la cual es una función real de la variable t, que alcanza un máximo en $t = t_0$. Entonces $(f \circ \alpha)'(t_0) = 0$. Esta derivada es, según la regla de la cadena, si $\alpha(t) = (x(t), y(t), z(t))$, la siguiente

$$\begin{aligned}(f \circ \alpha)'(t) &= \frac{\partial f}{\partial x}(\alpha(t))x'(t) + \frac{\partial f}{\partial y}(\alpha(t))y'(t) + \frac{\partial f}{\partial z}(\alpha(t))z'(t) \\ &= \operatorname{grad} f(\alpha(t)) \cdot \alpha'(t)\end{aligned}$$

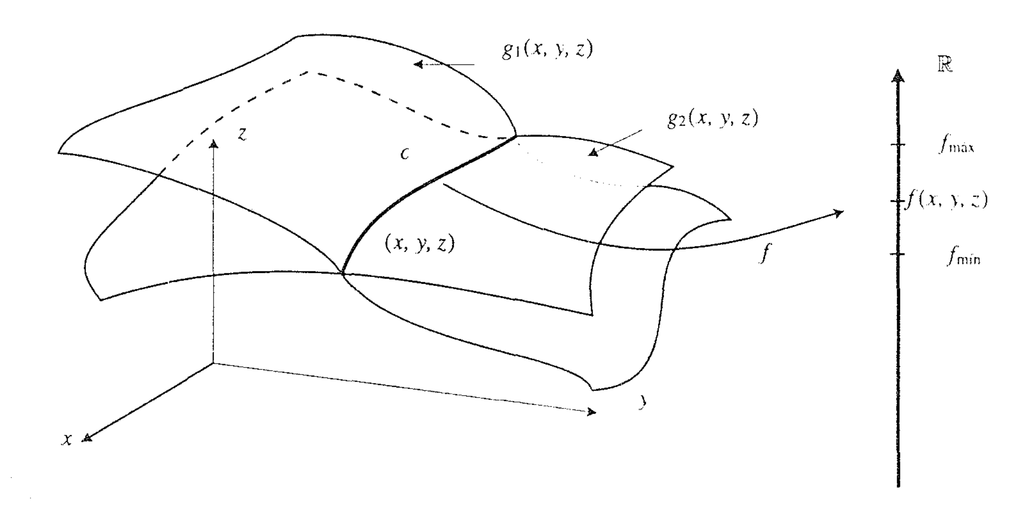


Figura 9. Gráfica de la intersección de las superficies g_1 y g_2.

donde $\alpha'(t) = (x'(t), y'(t), z'(t))$ es el vector derivada de α, que, como veremos en el próximo capítulo, es tangente a la curva C. Entonces

$$0 = (f \circ \alpha)'(t_0) = \text{grad } f(\alpha(t_0)) \cdot \alpha'(t_0) = \text{grad } f(p) \cdot \alpha'(t_0)$$

Siendo el vector $\alpha'(t_0)$ tangente a la curva C en el punto $\mathbf{p}$, y siendo grad $f(\mathbf{p}) \cdot \alpha'(t_0) = 0$, concluimos que el vector grad $f(\mathbf{p})$ debe ser normal a la curva C en $\mathbf{p}$. Es decir, grad $f(\mathbf{p})$ es un vector que se encuentra en el plano normal a la curva C en $\mathbf{p}$ (el plano que es atravesado perpendicularmente por C en $\mathbf{p}$).

Por otra parte, las expresiones $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$ se pueden ver como superficies de nivel (correspondientes al nivel cero) de las funciones $u = g_1(x, y, z)$, $u = g_2(x, y, z)$ respectivamente. Siendo así, sabemos que grad $g_1(\mathbf{p})$ es un vector normal a $g_1(x, y, z) = 0$ en $\mathbf{p}$ y grad $g_2(\mathbf{p})$ es un vector normal a $g_2(x, y, z) = 0$. Por otro lado, grad $g_1(\mathbf{p})$ y grad $g_2(\mathbf{p})$ son vectores normales a la curva C en $\mathbf{p}$. Siendo grad $f(\mathbf{p})$ un vector normal a la curva C en $\mathbf{p}$, concluimos que los tres vectores grad $f(\mathbf{p})$, grad $g_1(\mathbf{p})$, grad $g_2(\mathbf{p})$ se encuentran en el mismo plano (el plano normal a C en $\mathbf{p}$). Son, por tanto, linealmente dependientes. Si los vectores grad $g_1(\mathbf{p})$ y grad $g_2(\mathbf{p})$ son linealmente independientes, podemos expresar grad $f(\mathbf{p})$ como combinación lineal de ellos. En tales condiciones, deben existir constantes λ_1 y λ_2 que

$$\text{grad } f(\mathbf{p}) = \lambda_1 \text{ grad } g_1(\mathbf{p}) + \lambda_2 \text{ grad } g_2(\mathbf{p})$$

Concluimos entonces que una condición necesaria para que la función $u = f(x, y, z)$ sujeta a las restricciones $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$ alcance un extremo en (x, y, z) (donde se supone la independencia lineal de los vectores grad g_1 y grad g_2), es que se den constantes $\lambda_1, \lambda_2 \in \mathbb{R}$ de modo que

$$\text{grad } f = \lambda_1 \text{ grad } g_1 + \lambda_2 \text{ grad } g_2$$

Esta expresión, junto con las dos $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$ nos proporciona un conjunto de 5 ecuaciones con 5 incógnitas $(x, y, z, t, \lambda_1, \lambda_2)$, a saber

$$\frac{\partial f}{\partial x} = \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2 \frac{\partial g_2}{\partial x}, \quad \frac{\partial f}{\partial y} = \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2 \frac{\partial g_2}{\partial y}$$

$$\frac{\partial f}{\partial z} = \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2 \frac{\partial g_2}{\partial z}, \quad g_1(x, y, z) = 0, \quad g_2(x, y, z) = 0$$

Supongamos, por ejemplo, que la función $u = f(x, y, z) = xy + xz$, describe la temperatura en el punto (x, y, z). Consideremos la curva C intersección de las dos superficies $g_1(x, y, z) = x^2+y^2-1 = 0$ y $g_2(x, y, z) = y - z = 0$. Nos interesa saber cuál de los puntos de C está más caliente y cuál más frío.

Se trata entonces de obtener los extremos condicionados de la función $f(x, y, z) = xy+xz$, sujeta a las dos restricciones $g_1(x, y, z) = x^2 + y^2 - 1 = 0$ y $g_2(x, y, z) = y - z = 0$. Observe que la curva C es la intersección del cilindro $x^2 + y^2 = 1$ con el plano $z = y$ (figura 10).

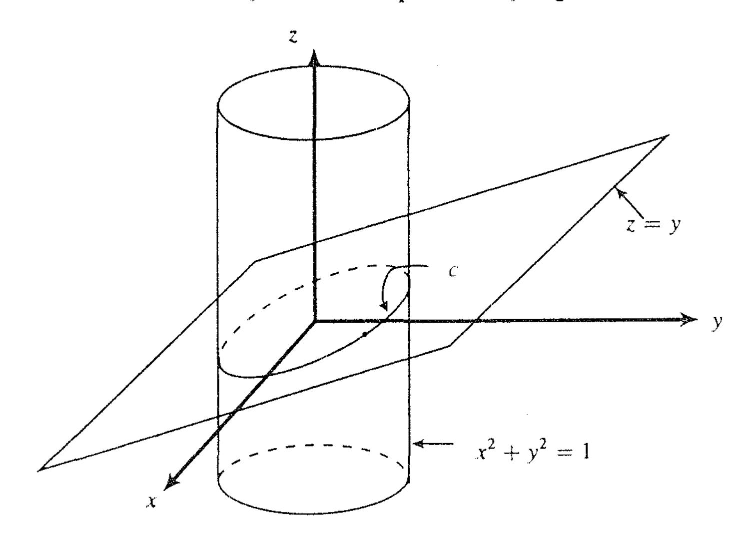


Figura 10. Gráfica de la curva C intersección del cilindro y el plano.

La expresión grad $f = \lambda_1$ grad $g_1 + \lambda_2$ grad g_2 se convierte en las tres ecuaciones

$$y + z = 2\lambda_1 x$$
$$x = 2\lambda_1 y + \lambda_2$$
$$x = -\lambda_2$$

de donde, eliminando λ_1 y λ_2 se obtiene $y^2 + zy - 2x^2 = 0$. Resolviendo esta expresión simultáneamente con $x^2 + y^2 = 1$ y $y = z$, obtenemos los cuatro puntos $\mathbf{p}_{1,2} = \left(\pm\frac{\sqrt{2}}{2}, \pm\frac{\sqrt{2}}{2}, \pm\frac{\sqrt{2}}{2}\right)$, $\mathbf{p}_{3,4} = \left(\pm\frac{\sqrt{2}}{2}, \mp\frac{\sqrt{2}}{2}, \mp\frac{\sqrt{2}}{2}\right)$. Se tiene $f(\mathbf{p}_{1,2}) = \left(\pm\frac{\sqrt{2}}{2}\right)^2 + \left(\pm\frac{\sqrt{2}}{2}\right)^2 = 1$, $f(\mathbf{p}_{3,4}) = \left(\pm\frac{\sqrt{2}}{2}\right)\left(\mp\frac{\sqrt{2}}{2}\right) + \left(\pm\frac{\sqrt{2}}{2}\right)\left(\mp\frac{\sqrt{2}}{2}\right) = -1$. Concluimos entonces que los puntos $\mathbf{p}_{1,2}$ son los más calientes de C, en tanto que los puntos $\mathbf{p}_{3,4}$ son los más fríos.

Vamos ahora a enunciar debidamente el teorema que establece condiciones necesarias para la existencia de extremos condicionados, en el caso general. Una vez entendidos los ejemplos presentados en la exposición previa, esperamos que para el lector sea natural el resultado que presentamos a continuación, cuya demostración queda fuera de los alcances de este texto.

Teorema 4.5.1 Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ definida en el conjunto abierto U de $\mathbb{R}^n$. Sean $g_1, g_2, \ldots, g_m: U \subseteq \mathbb{R}^n \to \mathbb{R}$, m funciones de clase $\mathscr{C}^1$ en U $(m < n)$. Sea

$$S = \{\mathbf{x} \in U | g_i(\mathbf{x}) = 0, i = 1, 2, \ldots, m\}$$

Sea $\mathbf{x}_0 \in S$ un punto de extremo condicionado de f, es decir tal que existe una bola $B \subset U$ con centro en $\mathbf{x}_0$ con la propiedad de que $f(\mathbf{x}_0) \leq f(\mathbf{x})$ ó $f(\mathbf{x}_0) \geq f(\mathbf{x})$ para toda $\mathbf{x}$ en $B \cap S$. Suponga que el determinante $\det\left(\frac{\partial g_i}{\partial x_j}(\mathbf{x}_0)\right) \neq 0$ para un conjunto de m variables x_j, tomadas del conjunto de n variables $\{x_1, x_2, \ldots, x_n\}$ de g_i. Entonces existen m números reales $\lambda_1, \lambda_2, \ldots, \lambda_m$ tales que se cumple

$$\operatorname{grad} f(\mathbf{x}_0) + \sum_{k=1}^{m} \lambda_k \operatorname{grad} g_k(\mathbf{x}_0) = 0$$ ■

A los números λ_k, $k = 1, 2, \ldots, m$ cuya existencia establece el teorema anterior se les llama MULTIPLICADORES DE LAGRANGE, y al método que el teorema proporciona para localizar los (posibles) extremos de la función f sujeta a las m restricciones $g_1(\mathbf{x}) = 0, \ldots, g_m(\mathbf{x}) = 0$, se le llama Método de los Multiplicadores de Lagrange. Para una demostración de este teorema puede consultarse, por ejemplo, la referencia [Ap], capítulo 7, sección 8

La expresión

$$\operatorname{grad} f(\mathbf{x}_0) + \sum_{k=1}^{m} \lambda_k \operatorname{grad} g_k(\mathbf{x}_0) = 0$$

que establece la dependencia lineal de los vectores $\operatorname{grad} f(\mathbf{x})$ y $\operatorname{grad} g_k(\mathbf{x})$, $k = 1, 2, \ldots, m$ en el punto $\mathbf{x}$, equivale a las n expresiones

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_j}(\mathbf{x}) = 0, \quad j = 1, 2, \ldots, n$$

que junto con las m ecuaciones $g_k(\mathbf{x}) = 0$, $k = 1, 2, \ldots, m$, producen un sistema de $m+n$ ecuaciones con $m+n$ incógnitas, a saber $x_1, x_2, \ldots, x_n$ (las coordenadas de $\mathbf{x}$) y $\lambda_1, \lambda_2, \ldots, \lambda_m$. Resolviendo este sistema localizamos a los candidatos $\mathbf{x} \in U$ a extremos condicionados de la función f.

Una manera concreta de obtener este sistema es formar primero la llamada Función de Lagrange

$$F(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(x_1, x_2, \ldots, x_n) + \sum_{k=1}^{m} \lambda_k g_k(x_1, x_2, \ldots, x_n)$$

Obsérvese que

$$\frac{\partial F}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_j}, \quad j = 1, 2, \ldots, n$$

$$\frac{\partial F}{\partial \lambda_k} = g_k \quad k = 1, 2, \ldots, m$$

de modo que las $m + n$ ecuaciones que se deben considerar son equivalentes a

$$\frac{\partial F}{\partial x_j} = 0, \quad \frac{\partial F}{\partial \lambda_k} = 0, \quad j = 1, 2, \ldots, n, \quad k = 1, 2, \ldots, m$$

lo que sugiere que se busca un punto crítico de la función de Lagrange F.

Por último, una vez localizados los puntos $\mathbf{x} \in U$ candidatos a extremos condicionados de la función f, se requerirá de un análisis adicional, que el mismo contexto del problema suele proporcionar de inmediato para concluir la naturaleza del extremo en cuestión (si es que éste existe). Sin embargo, un resultado que puede ser útil tener presente es aquél que nos dice que una función continua definida en un conjunto cerrado y acotado siempre alcanza el máximo y el mínimo en puntos de dicho conjunto (hecho que se expondrá en el apéndice de esta sección). Esto se aplica en nuestro caso cuando el conjunto S del teorema anterior, dado por las restricciones $g_k(\mathbf{x}) = 0$, $k = 1, 2, \ldots, n$ (que puede ser una curva, como en los dos ejemplos antes estudiados) es acotado —cerrado siempre, pues las funciones g_k que lo determinan son "buenas funciones"—. Así pues, si el conjunto S que marca la restricción de las variables de la función f, es acotado, podemos estar seguros que dentro de los valores que determinamos por el Método de los multiplicadores de Lagrange, se encontrará un máximo y un mínimo.

Ejemplo 1. Se quieren determinar los extremos de la función

$$f(x, y, z) = x^2 + y^2 + z^2$$

sujeta a la restricción

$$x^2 + \frac{1}{4}y^2 + \frac{1}{9}z^2 = 1$$

Formando la función de Lagrange

$$F(x, y, z, \lambda) = x^2 + y^2 + z^2 + \lambda(x^2 + \frac{1}{4}y^2 + \frac{1}{9}z^2 - 1)$$

Entonces el sistema a resolver es

$$\begin{aligned} \frac{\partial F}{\partial x} &= 2x + 2\lambda x = 0 \\ \frac{\partial F}{\partial y} &= 2y + \frac{1}{2}\lambda y = 0 \\ \frac{\partial F}{\partial z} &= 2z + \frac{2}{9}\lambda z = 0 \\ \frac{\partial F}{\partial \lambda} &= x^2 + \frac{1}{4}y^2 + \frac{1}{9}z^2 - 1 = 0 \end{aligned}$$

de donde se obtienen seis puntos que lo satisfacen, a saber $(\pm 1, 0, 0)$, $(0, \pm 2, 0)$, $(0, 0, \pm 3)$. Obsérvese que en este caso la superficie S que impone las restricciones es el elipsoide $x^2 + \frac{1}{4}y^2 + \frac{1}{9}z^2 = 1$, (una superficie acotada). Entonces entre los puntos encontrados deberá haber un máximo y un mínimo. Como $f(\pm 1, 0, 0) = 1$, $f(0, \pm 2, 0) = 4$, $f(0, 0, \pm 3) = 9$, se tiene que el mínimo (igual a 1) se logra en los puntos $(\pm 1, 0, 0)$ y el máximo (que vale 9) en los puntos $(0, 0, \pm 3)$. Es interesante darse cuenta de la geometría que hay detrás de este problema. La función $f(x, y, z) = x^2 + y^2 + z^2$ mide (el

cuadrado de) la distancia del punto (x, y, z) al origen. La superficie S de la restricción es un elipsoide con centro en el origen. Lo que hicimos entonces fue obtener los puntos del elipsoide que estaban más cerca (los puntos $(\pm 1, 0, 0)$) y más lejos (los puntos $(0, 0, \pm 3)$) del origen. ∎

Ejemplo 2. Se quieren hallar los extremos de la función $f(x, y, z) = xyz$ sujeta a las restricciones

$$g_1(x, y, z) = x^2 + y^2 + z^2 - 1 = 0 \quad \text{y} \quad g_2(x, y, z) = x + y + z = 0$$

Formamos la función de Lagrange

$$F(x, y, z, \lambda_1, \lambda_2) = \text{xyz} + \lambda_1(x^2 + y^2 + z^2 - 1) + \lambda_2(x + y + z)$$

y consideramos entonces el sistema

$$\frac{\partial F}{\partial x} = yz + 2\lambda_1 x + \lambda_2 = 0 \tag{1}$$

$$\frac{\partial F}{\partial y} = xz + 2\lambda_1 y + \lambda_2 = 0 \tag{2}$$

$$\frac{\partial F}{\partial z} = xy + 2\lambda_1 z + \lambda_2 = 0 \tag{3}$$

$$\frac{\partial F}{\partial \lambda_1} = x^2 + y^2 + z^2 - 1 = 0 \tag{4}$$

$$\frac{\partial f}{\partial \lambda_2} = x + y + z = 0 \tag{5}$$

De (2) y (3) se obtiene $(z - y)(x - 2\lambda_1) = 0$. Surgen dos opciones: a) $z = y$; b) $x = 2\lambda_1$. Si $z = y$, de (4) y (5) se obtiene el sistema $x^2 + 2y^2 = 1$, $x + 2y = 0$, que posee dos soluciones $x = \mp\frac{2}{\sqrt{6}}$, $y = \pm\frac{1}{\sqrt{6}}$. Así, en este caso ($z = y$) hay dos soluciones a nuestro sistema original, a saber $\mathbf{p}_1 = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$, $\mathbf{p}_2 = \left(\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)$. Regresando a la opción b), si $x = 2\lambda_1$, de (1) y (2) se obtiene que $(y - x)(z - x) = 0$. De donde o $y = x$ o $z = x$. En el primer caso nuevamente las ecuaciones (4) y (5) nos dan otras dos soluciones $\mathbf{p}_3 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$, $\mathbf{p}_4 = \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)$ y del segundo obtenemos $\mathbf{p}_5 = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$, $\mathbf{p}_6 = \left(\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)$. De nuevo observamos que el conjunto S que impone las restricciones al dominio f es la intersección de la esfera $x^2 + y^2 + z^2 = 1$ con el plano $x+y+z = 1$. Entonces es un círculo, el cual es acotado. Debe entonces haber un máximo y un mínimo entre los puntos localizados. Observando que $f(\mathbf{p}_1) = f(\mathbf{p}_3) = f(\mathbf{p}_5) = -\frac{1}{3\sqrt{6}}$, $f(\mathbf{p}_2) = f(\mathbf{p}_4) = f(\mathbf{p}_6) = \frac{1}{3\sqrt{6}}$, concluimos que la función tiene máximos en $\mathbf{p}_2, \mathbf{p}_4, \mathbf{p}_6$ (iguales a $\frac{1}{3\sqrt{6}}$) y mínimos en $\mathbf{p}_1, \mathbf{p}_3, \mathbf{p}_5$ (iguales a $-\frac{1}{3\sqrt{6}}$). ∎

Ejemplo 3. Retomemos el ejemplo 5 de la sección anterior. Se desea obtener las dimensiones de una caja que, con un volumen V dado, tuviera la menor superficie posible. Siendo x, y, z las dimensiones de la caja, la función de la cual queremos obtener el mínimo es $f(x, y, z) = 2xz + 2yz + xy =$ superficie de la caja. Notamos sin embargo, que las dimensiones x, y, z no pueden variar a su antojo: se debe tener $xyz = V$. Así, el problema consiste en localizar el mínimo de la funcióm f sujeta a la restricción $g(x, y, z) = xyz - V = 0$.

Formando la función de Lagrange

$$F(x, y, z, \lambda) = 2xz + 2yz + xy + \lambda(xyz - V)$$

obtenemos de ella el sistema

$$\frac{\partial F}{\partial x} = 2z + y + \lambda yz = 0 \qquad (1)$$
$$\frac{\partial F}{\partial y} = 2z + x + \lambda xz = 0 \qquad (2)$$
$$\frac{\partial F}{\partial z} = 2x + 2y + \lambda xy = 0 \qquad (3)$$
$$\frac{\partial F}{\partial \lambda} = xyz - V = 0 \qquad (4)$$

De (1) y (2) se obtiene $(y - x)(\lambda z + 1) = 0$, de donde $y = x$ (el factor $\lambda z + 1$ es no nulo, pues en caso contrario las mismas ecuaciones (1) y (2) nos conducirían a $z = 0$, valor imposible para z, pues no podrá satisfacer la ecuación (4)). De (3), con $y = x$, se obtiene $\lambda = -\frac{4}{x}$, y entonces de (2), se obtiene $z = \frac{x}{2}$. Así pues, de (4) llegamos a $x(x)\left(\frac{x}{2}\right) = V$, de donde $x = \sqrt[3]{2V} = y$, $z = \frac{1}{2}\sqrt[3]{2V}$. El contexto físico del problema nos dice que en este punto la función f debe tener un mínimo que vale $f\left(\sqrt[3]{2V}, \sqrt[3]{2V}, \frac{1}{2}\sqrt[3]{2V}\right) = 3(2V)^{2/3}$. Este es el mismo resultado al que llegamos de otro modo en el ejemplo 5 de la sección anterior (con $V = 100$). ∎

Ejemplo 4. Sea $\mathbf{p} = (x_0, y_0, z_0)$ un punto dado de $\mathbb{R}^3$ y $Ax + By + Cz = D$ un plano Π. Se quiere hallar la distancia del punto $\mathbf{p}$ al plano Π. Esta es la menor distancia que existe entre el punto $\mathbf{p}$ y los puntos del plano. Sabemos que la distancia entre $\mathbf{p}$ y (x, y, z) es resultado de

$$d = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

Se quiere hallar entonces el mínimo de esta función sujeta a la restricción $Ax + By + Cz = D$. Es decir, se quiere hallar el mínimo de la distancia entre $\mathbf{p} = (x_0, y_0, z_0)$ y el punto (x, y, z) cuando este punto pertenece al plano $Ax + By + Cz = D$. Por razones de simplicidad en los cálculos, en lugar de la función anterior podemos considerar mejor la función

$$f(x, y, z) = d^2 = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$$

Es claro que f tendrá un mínimo en un punto si y solamente si d lo tiene.

Tomemos entonces la función de Lagrange

$$F(x, y, z, \lambda) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 + \lambda(D - Ax - By - Cz)$$

Tenemos

$$\frac{\partial F}{\partial x} = 2(x - x_0) - \lambda A = 0 \qquad (1)$$
$$\frac{\partial F}{\partial y} = 2(y - y_0) - \lambda B = 0 \qquad (2)$$
$$\frac{\partial F}{\partial z} = 2(z - z_0) - \lambda C = 0 \qquad (3)$$
$$\frac{\partial F}{\partial \lambda} = D - Ax - By - Cz = 0 \qquad (4)$$

De (1), (2) y (3) (suponiendo por el momento que A, B y C no son cero) obtenemos

$$\frac{x - x_0}{A} = \frac{y - y_0}{B} = \frac{z - z_0}{C}$$

de modo que, según (4) se llega a (dejando todo en términos de z)

$$D - A\left(\frac{A}{C}(z - z_0) + x_0\right) - B\left(\frac{B}{C}(z - z_0) + y_0\right) - Cz = 0$$

de donde

$$z - z_0 = \frac{C(D - Ax_0 - By_0 - Cz_0)}{A^2 + B^2 + C^2}$$

como $x - x_0 = \frac{A}{C}(z - z_0)$, $y - y_0 = \frac{B}{C}(z - z_0)$, obtenemos también

$$x - x_0 = \frac{A(D - Ax_0 - By_0 - Cz_0)}{A^2 + B^2 + C^2}, \quad y - y_0 = \frac{B(D - Ax_0 - By_0 - Cz_0)}{A^2 + B^2 + C^2}$$

Al despejar x, y, z obtenemos la solución al problema, pues siendo ésta la única reportada por el método de los multiplicadores de Lagrange, es claro que se debe tratar del mínimo procurado para la función f. En tal punto se tiene

$$f(x, y, z) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = \frac{(D - Ax_0 - By_0 - Cz_0)^2}{A^2 + B^2 + C^2}$$

de modo que la distancia mínima buscada es $d = \sqrt{f(x, y, z)}$

$$d = \frac{|Ax_0 + By_0 + Cz_0 - D|}{\sqrt{A^2 + B^2 + C^2}}$$

Esta es, efectivamente, la fórmula de la distancia entre el punto $\mathbf{p} = (x_0, y_0, z_0)$ y el plano $Ax + By + Cz = D$. ■

Ejemplo 5. Una desigualdad muy importante en matemáticas, conocida como *desigualdad de Hölder*, establece que si $a_1, \ldots, a_n$, $x_1, \ldots, x_n$ son números reales no negativos, y los números positivos p, q son exponentes conjugados (lo cual significa que $p > 1$ y $\frac{1}{p} + \frac{1}{q} = 1$), entonces

$$\sum_{i=1}^{n} a_i x_i \leq \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} x_i^q\right)^{1/q}$$

Veamos cómo usar el método de los multiplicadores de Lagrange para establecer la validez de esta desigualdad. Consideremos la función f de las variables $x_1, \ldots, x_n$ dada por

$$f(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} x_i^q\right)^{1/q}$$

Por ahorro en la notación llamemos $K = \left(\sum_{i=1}^{n} a_i^p\right)^{1/p}$. Se desea encontrar el mínimo de esta función sujeta a la restricción

$$A = \sum_{i=1}^{n} a_i x_i$$

(A es un número dado).

Formamos entonces la función de Lagrange

$$F(x_1, \ldots, x_n, \lambda) = K\left(\sum_{i=1}^{n} x_i^q\right)^{1/q} + \lambda\left(\sum_{i=1}^{n} a_i x_i - A\right)$$

de donde

$$\frac{\partial F}{\partial x_j} = K\frac{1}{q}\left(\sum_{i=1}^{n} x_i^q\right)^{1/q-1} q x_j^{q-1} + \lambda a_j = 0, \qquad j = 1, 2, \ldots, n$$

$$\frac{\partial F}{\partial \lambda} = \sum_{i=1}^{n} a_i x_i - A = 0$$

Tenemos así el sistema de $n+1$ ecuaciones con $n+1$ indeterminadas $(x_1, x_2, \ldots, x_n, \lambda)$ siguiente

$$K\left(\sum_{i=1}^{n} x_i^q\right)^{1/q-1} x_j^{q-1} + \lambda a_j = 0, \qquad j = 1, 2, \ldots, n$$

$$\sum_{i=1}^{n} a_i x_i - A = 0$$

De la primera y la j-ésima ecuación, $2 \leq j \leq n$, obtenemos, despejando λ de ellas e igualando, que (suponiendo, sin perder la generalidad, que $a_1 \neq 0$ y $a_j \neq 0$)

$$-\frac{1}{a_j}K\left(\sum_{i=1}^{n} x_i^q\right)^{1/q-1} x_j^{q-1} = -\frac{1}{a_1}K\left(\sum_{i=1}^{n} x_i^q\right)^{1/q-1} x_1^{q-1}$$

de donde se obtiene que

$$x_j = \left(\frac{a_j}{a_1}\right)^{1/(q-1)} x_1, \qquad j = 1, 2, \ldots, n$$

De la última ecuación obtenemos, sustituyendo estos valores de x_j

$$A = a_1 x_1 + \sum_{j=2}^{n} a_j x_j = a_1 x_1 + \sum_{j=2}^{n} a_j \left(\frac{a_j}{a_1}\right)^{1/(q-1)} x_1$$

$$= \frac{x_1}{a_1^{1/(q-1)}} \sum_{j=1}^{n} a_j^{q/(q-1)}$$

de donde, en virtud de que $\frac{1}{p} + \frac{1}{q} = 1$

$$x_1 = \frac{a_1^{1/(q-1)} A}{\sum_{j=1}^{n} a_j^p} = \frac{A}{K^p} a_1^{1/(q-1)}$$

y entonces

$$x_j = \left(\frac{a_j}{a_1}\right)^{1/(q-1)} x_1 = \frac{A}{K^p} a_j^{1/(q-1)}, \quad j = 1, 2, \ldots, n$$

Tenemos pues que el método de los multiplicadores de Lagrange nos reporta un único punto de solución

$$x_j = \frac{A}{K^p} a_j^{1/(q-1)}, \quad j = 1, 2, \ldots, n$$

Invitamos al lector a que considere el aspecto geométrico de nuestro problema en los casos $n = 2$ y $n = 3$, para que se convenza de que esta solución debe corresponder de hecho a un mínimo de la función f, que vale

$$\begin{aligned} f(x_1, \ldots, x_n) &= K \left(\sum_{j=1}^{n} x_j^q\right)^{1/q} = K \left(\sum_{j=1}^{n} \frac{A^q}{K^{pq}} a_j^{q/(q-1)}\right)^{1/q} \\ &= \frac{KA}{K^p} \left(\sum_{j=1}^{n} a_j^p\right)^{1/q} = \frac{KA}{K^p} K^{p/q} = AK^{p/q+1-p} \\ &= A \end{aligned}$$

de modo que, siendo A el mínimo condicionado de f se tiene

$$A \leq f(x_1, \ldots, x_n)$$

para toda $(x_1, \ldots, x_n)$ tal que $\sum_{i=1}^{n} a_i x_i = A$. Es decir,

$$\sum_{i=1}^{n} a_i x_i \leq \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} x_i^q\right)^{1/q}$$

como se quería demostrar. ■

Ejemplo 6. (En un tono menos serio: un problema práctico de economía culinaria). Cuando a doña Gela se le ocurrió hacer una sencilla gelatina usando solamente un poco de grenetina, azúcar y licor de plátano que unos compadres producían en una industria cerca de su pueblo, nunca imaginó que estaba descubriendo la receta que sería la fuente de ingresos con los cuales iba a poder sostener a su familia por largos años. Durante mucho tiempo doña Gela estuvo surtiendo pequeños pedidos para fiestas: un promedio de cinco gelatinas por semana era su producción estándar. Doña Gela es una persona con un gran gusto por la matemática, así que en todo ese tiempo, además de obtener algunos ingresos extras, se dedicó también a establecer un modelo que le proporcionara el número de gelatinas que podía producir en términos de las cantidades de los ingredientes que usaba en su

elaboración. La fórmula a la que llegó es $g = x^a y^b z^c$, con a, b y c algunas constantes secretas que doña Gela guarda celosamente en su memoria, x la cantidad en kg de grenetina, y la cantidad en kg de azúcar, z la cantidad en litros de licor de plátano y g la cantidad de gelatinas que se pueden hacer con tales cantidades de ingredientes. Ahora le llegaba una oportunidad importante a Doña Gela: un pedido de cuando menos 100 gelatinas mensuales para una cadena de salones de fiestas infantiles. Ante esta nueva etapa de gran producción de gelatinas, doña Gela se planteó el problema de saber qué cantidades de ingredientes tendría que comprar mensualmente para producir una cantidad determinada de gelatinas, de modo que le representara el menor gasto posible en la compra de los ingredientes. Es decir, si N es el número de gelatinas que tiene que hacer, quería saber los valores de x, y y z para tal producción, y cuyo precio en conjunto (conociendo los precios de cada uno de ellos), representaran la menor inversión posible (en la compra de la materia prima).

Veamos cómo con el método de los multiplicadores de Lagrange estudiado en esta sección, es posible resolver el problema de la óptima producción de gelatinas de doña Gela. Sean A, B y C los precios de la grenetina, el azúcar y el licor de plátano, respectivamente (A y B en pesos por kilo y C en pesos por litro). Entonces la función

$$f(x, y, z) = Ax + By + Cz$$

nos da el costo invertido al comprar x kg de grenetina, y kg de azúcar y z litros de licor de plátano. Nos interesa obtener el mínimo de esta función, cuando se tiene la restricción de que el número de gelatinas producidas debe ser N; es decir, usando la fórmula obtenida por doña Gela; dicha restricción se ve

$$x^a y^b z^c = N$$

Consideremos entonces la función de Lagrange

$$F(x, y, z, \lambda) = Ax + By + Cz + \lambda(x^a y^b z^c - N)$$

El sistema a resolver es

$$\frac{\partial F}{\partial x} = A + \lambda a x^{a-1} y^b z^c = 0 \qquad (1)$$

$$\frac{\partial F}{\partial y} = B + \lambda b x^a y^{b-1} z^c = 0 \qquad (2)$$

$$\frac{\partial F}{\partial z} = C + \lambda c x^a y^b z^{c-1} = 0 \qquad (3)$$

$$\frac{\partial F}{\partial \lambda} = x^a y^b z^c - N = 0 \qquad (4)$$

de (1), (2) y (3) se obtiene que

$$y = \frac{bA}{aB}x, \quad z = \frac{cA}{aC}x$$

Sustituyendo en (4) se obtiene

$$x^a \left(\frac{bA}{aB}\right)^c x^b \left(\frac{cA}{aC}\right)^c x^c = N$$

de donde

$$\bar{x} = \left(\frac{bA}{aB}\right)^{-b}\left(\frac{cA}{aC}\right)^{-c} N^{\frac{1}{a+b+c}}$$

y entonces

$$\bar{y} = \left(\frac{bA}{aB}\right)^{-b+1}\left(\frac{cA}{aC}\right)^{-c} N^{\frac{1}{a+b+c}}, \quad \bar{z} = \left(\frac{bA}{aB}\right)^{-b}\left(\frac{cA}{aC}\right)^{-c+1} N^{\frac{1}{a+b+c}}$$

Siendo ésta la única solución reportada por el método de los multiplicadores de Lagrange, es claro que debe de tratarse de un mínimo para la función f. Así, para poder producir N gelatinas y gastar lo menos posible en la compra de los ingredientes, habrá que comprar $\bar{x}$ kg de grenetina, $\bar{y}$ kg de azúcar y $\bar{z}$ lt de licor de plátano. ■

Apéndice Extremos absolutos de funciones en regiones compactas

Uno de los resultados clásicos sobre funciones reales de una variable real, definidas en un intervalo cerrado y acotado $[a, b]$, donde la función es continua, es que ésta alcanza su máximo y mínimo absolutos dentro del intervalo $[a, b]$. Más aún, si la función $f: [a, b] \to \mathbb{R}$ es continua, entonces ocurren $x_0, x_1 \in [a, b]$, que $M = f(x_0) \geq f(x)\, \forall x \in [a, b], m = f(x_1) \leq f(x)\, \forall x \in [a, b]$. Al valor M (respectivamente, m) se le llama *máximo (mínimo) absoluto de la función f en el intervalo* $[a, b]$. Se sabe además que los puntos x_0, x_1 donde f alcanza sus extremos absolutos, son o los extremos del intervalo $[a, b]$ (es decir, los puntos $x = a$ o $x = b$), o bien, si es un punto del intervalo abierto (a, b), éste debe ser un punto crítico, es decir, un punto en donde la derivada f' es cero o no existe. Las gráficas de la figura 11 ilustran estos hechos.

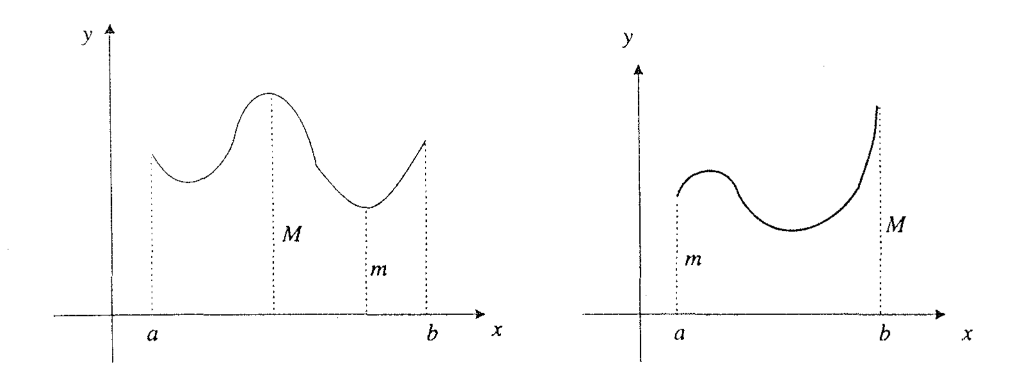


Figura 11. Extremos absolutos de una función continua $f: [a, b] \to \mathbb{R}$.

Este resultado sigue siendo cierto para funciones reales continuas de varias variables definidas en cierto tipo de regiones $K \subset \mathbb{R}^n$, con propiedades "análogas" (desde el punto de vista topológico) a las de los intervalos cerrados y acotados de $\mathbb{R}$. Estas regiones se llaman *compactas* y se caracterizan por ser conjuntos *cerrados* y *acotados* de $\mathbb{R}^n$, es decir, conjuntos cuyo complemento $\mathbb{R}^n - K$ es un conjunto abierto en $\mathbb{R}^n$, de modo que se puedan incluir en una bola B con centro en el origen y algún $r > 0$. Incluso, se tiene el resultado siguiente.

Teorema 4.5.2 Sea $f: K \subset \mathbb{R}^n \to \mathbb{R}$ una función real definida en el conjunto compacto K de $\mathbb{R}^n$. Supongamos que f sea continua. Entonces existen puntos $\mathbf{x}_0, \mathbf{x}_1 \in K$ tales que $M = f(\mathbf{x}_0) \geq f(\mathbf{x})\ \forall \mathbf{x} \in K, m = f(\mathbf{x}_1) \leq f(\mathbf{x})\ \forall \mathbf{x} \in K$. ■

Al valor M (a m) se le llama *máximo (mínimo) absoluto de f en K*. Entonces lo que establece el teorema anterior (cuya demostración se puede consultar por ejemplo en [Ap]), es que la función continua f definida en el conjunto compacto K alcanza su máximo y mínimo absolutos en puntos de K. Si además suponemos que la función f es diferenciable (por ejemplo en un conjunto abierto U que contenga a K), se puede demostrar que los extremos absolutos de f ocurren: 1) en la frontera de K; o bien, 2) en puntos interiores de K, donde las derivadas parciales de f se deben anular (es decir, si el (los) extremo(s) absoluto(s) de f se encuentran en puntos del interior de K, estos deben ser puntos críticos de f).

En el caso de funciones diferenciables de dos variables, $f: K \subset \mathbb{R}^2 \to \mathbb{R}$, definidas en un conjunto compacto K (la función f continua en K y diferenciable en un abierto U que contiene a K), limitado por una curva cerrada del plano, digamos que dada por la ecuación $g(x, y) = 0$, el problema de determinación de los extremos absolutos de f en K se limita a determinar los puntos críticos de *f dentro de K* (pues en ellos se podrían encontrar los extremos procurados), y luego, determinar los extremos de *f en los puntos de la curva dada $g(x, y) = 0$*, la cual es la frontera de K, donde también se pueden encontrar los extremos absolutos de f. Estos últimos cálculos (para determinar los extremos de f en la frontera de K) se pueden realizar resolviendo el problema de extremos condicionados de la función f con la restricción de que sus variables independientes x, y satisfagan la relación $g(x, y) = 0$. Comparando los valores de f en los puntos así encontrados, se puede decidir cuál de ellos es el máximo y cuál el mínimo absoluto de f en el conjunto compacto K.

Veamos algunos ejemplos.

Ejemplo 1. Se han de determinar los extremos absolutos de la función $f(x, y) = x^2 + 3y^2$, en la región $K = \{(x, y)|x^2 - 2x + y^2 - 3 \leq 0\}$. Obsérvese que la curva $g(x, y) = x^2 - 2x + y^2 - 3 = (x-1)^2 + y^2 - 4 = 0$, que limita a la región K, es un círculo con centro en $(1, 0)$ y radio 2. Se determinarán entonces los valores máximo y mínimo de $f(x, y)$ con $(x, y) \in K$. En principio, se localizan los puntos críticos de f que estén dentro de K. Resolviendo $\frac{\partial f}{\partial x} = 2x = 0$, $\frac{\partial f}{\partial y} = 6y = 0$, se obtiene el punto $\mathbf{p}_1 = (0, 0)$ dentro de K. Pasamos ahora a determinar los extremos de *f en la frontera de K*. Para ello, resolvemos el problema de extremos condicionados de f con la restricción $g(x, y) = 0$. Se tiene entonces la función de Lagrange

$$F(x, y) = x^2 + 3y^2 + \lambda(x^2 - 2x + y^2 - 3)$$

de donde

$$\frac{\partial F}{\partial x} = 2x + \lambda(2x - 2) = 0, \quad \frac{\partial F}{\partial y} = 6y + 2\lambda y = 0$$
$$\frac{\partial F}{\partial \lambda} = x^2 - 2x + y^2 - 3 = 0$$

de donde se obtienen los puntos $\mathbf{p}_2 = (-1, 0)$, $\mathbf{p}_3 = (3, 0)$. Se tiene entonces $f(\mathbf{p}_1) = f(0, 0) = 0$, $f(\mathbf{p}_2) = f(-1, 0) = 1$, $f(3, 0) = 9$. Así pues, el mínimo absoluto de f en K se encuentra en el punto $\mathbf{p}_1 = (0, 0)$ (dentro de K, en un punto crítico de f), el cual vale 0, y el máximo absoluto de f en K se encuentra en el punto $\mathbf{p}_2 = (3, 0)$ (en la frontera de K), el cual vale 9. Para que se convenza de los resultados obtenidos, invitamos al lector a imaginar la geometría del problema. ■

Ejemplo 2. Se han de determinar los extremos absolutos de la función $f(x, y) = x^2y^3(1 - x - y)$, en la región $K = \{(x, y) | |x| + |y| \leq 1\}$. Esta región corresponde a un cuadrado con vértices en los puntos $(\pm 1, 0)$ y $(0, \pm 1)$, el cual está limitado por las rectas: L_1: $y = 1 - x$, $0 \leq x \leq 1$, L_2: $y = 1 + x$, $-1 \leq x \leq 0$, L_3: $y = -1 - x$, $-1 \leq x \leq 0$, L_4: $y = -1 + x$, $0 \leq x \leq 1$ (ver figura 12).

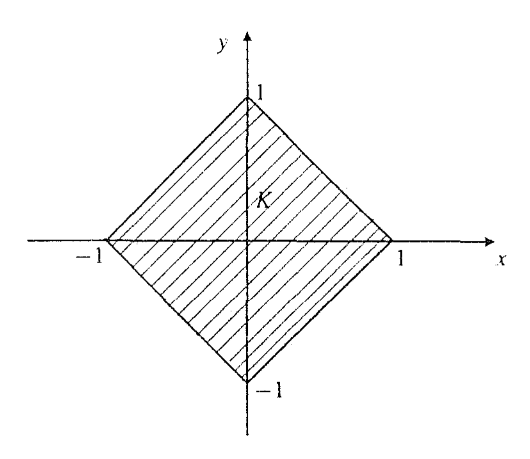


Figura 12. Región K del ejemplo 2.

Primero se localizan los puntos críticos de f en K. Se tiene

$$\frac{\partial f}{\partial x} = xy^3(2 - 3x - 2y) = 0, \quad \frac{\partial f}{\partial y} = x^2y^2(3 - 3x - 4y) = 0$$

de donde se obtienen los puntos $\mathbf{p}_1 = (1/3, 1/2)$, $\mathbf{p}_2 = (0, t)$, $\mathbf{p}_3 = (t, 0)$, $-1 < t < 1$, que se encuentran en el interior de K. Obtengamos ahora los extremos de f en la frontera de K. Sobre L_1 la función f se ve como $\phi_1(x) = f(x, 1 - x) = x^2(1 - x)^3(1 - x - (1 - x)) = 0$. Sobre L_2, f se ve como $\phi_2 = f(x, 1 + x) = x^2(1 + x)^3(1 - x - (1 + x)) = -2x^3(1 + x)^3$, con $-1 \leq x \leq 0$. Determinamos los extremos absolutos de $\phi_2(x)$ en el intervalo $[-1, 0]$, encontrando que se tienen mínimos absolutos en $x = -1$, $x = 0$, donde ϕ_2 vale 0, y el máximo absoluto en $x = -1/2$, donde ϕ_2 vale 1/32. Sobre L_2 se tiene entonces que f alcanza su mínimo absoluto en los puntos $(0, 1)$ y $(-1, 0)$ donde f vale 0, y el máximo absoluto en $(-1/2, 1/2)$ donde f vale 1/32. Para L_3, f se ve como $\phi_3(x) = f(x, -1 - x) = x^2(-1 - x)^3(1 - x - (-1 - x)) = -2x^2(1 + x)^3$, con $-1 \leq x \leq 0$. Los extremos absolutos de $\phi_3(x)$ en $[-1, 0]$ se encuentran en $x = -1$, $x = 0$, mínimos absolutos donde ϕ_3 vale 0, y $x = -2/5$, máximo absoluto en que ϕ_3 vale $-216/3125$. Es decir, para f en L_3 se tienen mínimos absolutos en los puntos $(0, -1)$ y $(-1, 0)$ donde f vale 0, y máximo absoluto en el punto $(-2/5, -3/5)$ donde f vale $-216/3125$. Finalmente, para f en L_4, se tiene $\phi_4(x) = f(x, -1 + x) = x^2(-1 + x)^3(1 - x - (-1 + x)) = -2x^2(x - 1)^4$, con $0 \leq x \leq 1$. Los extremos absolutos de $\phi_4(x)$ en $[0, 1]$ se encuentran en $x = 0$, $x = 1$, mínimos absolutos, donde ϕ_4 vale 0, y $x = 1/3$, máximo absoluto, donde ϕ_4 vale $-32/729$. Se tiene entonces que f en L_4 alcanza su mínimo absoluto en los puntos $(0, -1)$ y $(1, 0)$ donde f vale 0, y su máximo absoluto en $(1/3, -2/3)$, en que f vale $-32/729$. En resumen, la función f alcanza, sobre la frontera de la región K, su mínimo absoluto en los puntos $(x, 1 - x)$, $0 \leq x \leq 1$, $(-1, 0)$ y $(0, -1)$ donde f vale 0, y su máximo absoluto en el punto $(-1/2, 1/2)$ donde f vale 1/32. Comparando con los valores de f en los puntos críticos que se hallaron dentro de K, $f(t, 0) = f(0, t) = 0$, $f(1/3, 1/2) = 1/432$,

vemos que los extremos absolutos de f en K son: el mínimo absoluto en $(-2/5, -3/5)$ donde f vale $-216/3125$, y el máximo absoluto en el punto $(-1/2, 1/2)$ donde f vale $1/32$. ■

Las ideas anteriores, establecidas para funciones continuas de dos variables $f(x, y)$ definidas en conjuntos compactos K de $\mathbb{R}^2$, se extienden de manera natural para funciones continuas de n variables $f(x_1, x_2, \ldots, x_n)$ definidas en conjuntos compactos K de $\mathbb{R}^n$.

Ejemplo 3. Se han de determinar los extremos absolutos de la función $f(x, y, z) = xy^2z^3$ en la región $K = \{(x, y, z)|x^2 + y^2 + z^2 \leq 1\}$ = bola unitaria cerrada. Los puntos críticos de f en K se hallan en el sistema

$$\frac{\partial f}{\partial x} = y^2z^3 = 0, \quad \frac{\partial f}{\partial y} = 2xyz^3 = 0, \quad \frac{\partial f}{\partial z} = 3xy^2z^2 = 0$$

de donde se obtiene el único punto crítico $(0, 0, 0)$, el cual se encuentra en el interior de K. Para investigar los extremos de f en la frontera de K, consideramos el problema de extremos condicionados de f con la restricción $g(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$. La función de Lagrange es

$$F(x, y, z, \lambda) = xy^2z^3 + \lambda(x^2 + y^2 + z^2 - 1)$$

de donde se obtiene el sistema

$$\begin{aligned} \frac{\partial F}{\partial x} &= y^2z^3 + 2\lambda x = 0, \\ \frac{\partial F}{\partial y} &= 2xyz^3 + 2\lambda y = 0 \\ \frac{\partial F}{\partial z} &= 3xy^2z^2 + 2\lambda z = 0, \\ \frac{\partial F}{\partial \lambda} &= x^2 + y^2 + z^2 - 1 = 0 \end{aligned}$$

De las primeras tres ecuaciones se tiene (igualando λ)

$$\frac{yz}{2x} = \frac{xz}{y} = \frac{3xy}{2z}$$

de donde se obtiene que $(y^2 - 2x^2)z = 0$, $2z^2 = 3y^2$, expresiones que junto con la restricción dada, producen los siguientes puntos como candidatos a extremos condicionados: $\mathbf{p}_{1\,2} = (\pm 1, 0, 0)$, $\mathbf{p}_{3\,4\,5\,6\,7\,8\,9\,10} = (\pm 1/\sqrt{6}, \pm 1/\sqrt{3}, \pm 1/\sqrt{2})$. Se tiene entonces $f(0, 0, 0) = 0$,

$$f(1/\sqrt{6}, \pm 1/\sqrt{3}, 1/\sqrt{2}) = f(-1/\sqrt{6}, \pm 1/\sqrt{3}, -1/\sqrt{2}) = \frac{1}{12\sqrt{3}}$$
$$f(1/\sqrt{6}, \pm 1/\sqrt{3}, -1/\sqrt{2}) = f(-1/\sqrt{6}, \pm 1/\sqrt{3}, 1/\sqrt{2}) = -\frac{1}{12\sqrt{3}}$$

de donde se ve que los extremos absolutos de f se encuentran en la frontera de K (en puntos (x, y, z) donde $x^2 + y^2 + z^2 = 1$), estando los máximos absolutos en los cuatro puntos $(1/\sqrt{6}, \pm 1/\sqrt{3}, 1/\sqrt{2})$, $(-1/\sqrt{6}, \pm 1/\sqrt{3}, -1/\sqrt{2})$, en que f vale $\frac{1}{12\sqrt{3}}$, y los mínimos absolutos en los cuatro puntos $(1/\sqrt{6}, \pm 1/\sqrt{3}, -1/\sqrt{2})$, $(-1/\sqrt{6}, \pm 1/\sqrt{3}, 1/\sqrt{2})$, donde f vale $-\frac{1}{12\sqrt{3}}$. ■

Ejercicios (Capítulo 4, Sección 5)

1. Determine los extremos de la función $f(x, y) = 2x - y$, con la restricción $g(x, y) = 3x^2 + 2y^2 - 33/2 = 0$.
2. Determine los extremos de la función $f(x, y) = x + 3y$, con la restricción $g(x, y) = 2x^2 + y^2 - 38 = 0$.
3. Determine los extremos de la función $f(x, y) = 3x - 7y$, con la restricción $g(x, y) = 4x^2 + 3y^2 - 892/3 = 0$.
4. Halle los extremos de la función $f(x, y) = -2x - 5y$, sujeta a la restricción $g(x, y) = 3x^2 + y^2 - 237 = 0$.
5. Halle los extremos de la función $f(x, y) = x+2y$, sujeta a la restricción $g(x, y) = x^2+y^2-5 = 0$.
6. Determine el valor máximo que puede tomar el producto de dos números positivos, si la suma de éstos debe ser 20.
7. Determine el valor máximo que puede tomar el producto de tres números positivos, si la suma de éstos debe ser 20.
8. Determine los extremos de la función $f(x, y) = x^2y^2$, si (x, y) se encuentra en el círculo unitario $x^2 + y^2 = 1$.
9. Halle los extremos de la función $f(x, y) = x^2y$, si (x, y) se encuentra en la elipse $2x^2 + y^2 = 3$.
10. Determine los extremos de la función $f(x, y) = x^2 + 8y$, si (x, y) se encuentra en la elipse $x^2 + 4y^2 = 5$.

En los ejercicios 11–15 se da una función de dos variables $z = f(x, y)$, de la que se han de determinar los extremos cuando sus variables están sujetas a la restricción $g(x, y) = 0$. Estos se pueden resolver como problemas de extremos condicionados, o bien haciendo explícita una de las variables x o y de la restricción dada y sustituyendo ésta en la función f dada, convirtiendo así el problema en la determinación de extremos locales de una función de una sola variable. Resuelva estos ejercicios de las dos maneras mencionadas.

11. $f(x, y) = x^3 + y^3 + 3x^2y$, $g(x, y) = x + y - 3 = 0$.
12. $f(x, y) = x^2 + 2y^2 + 3xy + 1$, $g(x, y) = x - y - 1 = 0$.
13. $f(x, y) = 2x^3 + x^2y$, $g(x, y) = x + y - 2 = 0$.
14. $f(x, y) = 3x^3 + y^3 + x^2y$, $g(x, y) = x + y - 5 = 0$.
15. $f(x, y) = 16x^3 + y^3 + 3x^2y$, $g(x, y) = x + y - 6$.
16. Determine los extremos de la función $f(x, y, z) = 2x + 3y + z$, sujeta a la restricción $g(x, y, z) = 4x^2 + 3y^2 + z^2 - 80 = 0$.
17. Determine los extremos de la función $f(x, y, z) = -x - 4y + 5z$, sujeta a la restricción $g(x, y, z) = 5x^2 + y^2 + z^2 - 1030 = 0$.
18. Halle los extremos de la función $f(x, y, z) = x - y - z$, si (x, y, z) está en el elipsoide $x^2 + 2y^2 + z^2 = 50$.

19. Halle los extremos de la función $f(x, y, z) = 2x - 3y + z$, si (x, y, z) está en el elipsoide $4x^2 + y^2 + z^2 = 704$.

20. Determine los puntos (x, y, z) del elipsoide $2x^2 + 4y^2 + 5z^2 = 70$ de modo que la suma de su primera y tercera coordenadas sea la mayor y la menor posible.

21. Demuestre que la función $f(x, y, z) = x^3 + x^2y + 0.5xy^2 + xz$, cuando (x, y, z) satisface $x + y + z + 1 = 0$, tiene un máximo en el punto $\mathbf{p}_1 = (-1, 2, -2)$ y un mínimo en el punto $\mathbf{p}_2 = (1, 0, -2)$. Resuelva este problema: **a.** como problema de extremos condicionados; **b.** despejando una de las variables de la restricción y sustituyéndola en la función f, para resolver entonces el problema de determinación de extremos locales de la función de dos variables obtenida.

22. Demuestre que la función $f(x, y, z) = x^3 + x^2y + xy^2 + xz$, cuando (x, y, z) satisface $x+y+z+1 = 0$, tiene un máximo en el punto $\mathbf{p}_1 = (-5/9, 7/9, -11/9)$ y un mínimo en el punto $\mathbf{p}_2 = (1, 0, -2)$. Resuelva este problema: **a.** como problema de extremos condicionados; **b.** despejando una de las variables de la restricción y sustituyéndola en la función f, para resolver el problema de determinación de extremos locales de la función de dos variables obtenida.

23. Considere la función $f(x, y, z) = x^3 + 2y^2 + 2z^2 + x^2y + xy^2 + xz$, con la restricción $g(x, y, z) = x + y + z - 1 = 0$. Siga el método de los multiplicadores de Lagrange presentado en esta sección, para obtener que los puntos $\mathbf{p}_1 = (5/9, 2/9, 2/9)$, $\mathbf{p}_2 = (-1, 1, 1)$ son los candidatos a extremos condicionados de la función dada. Convierta este problema en el de determinación de extremos locales de una función de dos variables, para verificar que la función tiene un mínimo en $\mathbf{p}_1$, y que *no tiene extremo* en $\mathbf{p}_2$. Verifique que $f(-1, 1, 1) = 2$. Calcule $f(-0.9, 1.1, 0.8)$ y $f(-1.1, 1, 1.1)$. ¿Le ayudan estos resultados a confirmar que f no tiene extremo condicionado en $\mathbf{p}_2$?

24. Considere la función $f(x, y, z) = x^3 + y^2 + z^2 + x^2y + xy^2 + xz$, con la restricción $g(x, y, z) = x + y + z - 1 = 0$. Siga el método de los multiplicadores de Lagrange expuesto en esta sección para obtener que los puntos $\mathbf{p}_1 = (1/3, 1/3, 1/3)$, $\mathbf{p}_2 = (-1/3, 2/3, 2/3)$ son los candidatos a extremos condicionados de la función dada. Convierta este problema en el de determinación de extremos locales de una función de dos variables, para verificar que la función tiene un mínimo en $\mathbf{p}_1$, y que *no tiene extremo* en $\mathbf{p}_2$. Localice puntos (x, y, z), tales que $x + y + z = 1$, cercanos de $\mathbf{p}_2$ para los cuales $f(x, y, z) > f(\mathbf{p}_2)$ y $f(x, y, z) < f(\mathbf{p}_2)$.

25. Siga el método de los multiplicadores de la Lagrange para determinar los semiejes de la elipse $5x^2 - 6xy + 5y^2 - 32 = 0$.

26. Calcule el área de la elipse $25x^2 - 14xy + 25y^2 - 288 = 0$. (Sugerencia: el área de una elipse con semiejes a y b es igual a πab; determine a y b con el método de los multiplicadores de Lagrange).

27. Determine los puntos más cercanos y más alejados del origen de la curva cerrada $x^2+y^2+xy = 4$.

28. Determine los puntos más cercanos y más alejados del origen de la curva $(x-1)^2+(y-1)^2 = 4$. Resuelva este problema: **a.** con el método de los multiplicadores de Lagrange; **b.** con ideas de geometría analítica.

29. Determine los puntos más cercanos y más alejados del origen de la superficie $(x - 2)^2 + (y - 1)^2 + (z - 2)^2 = 16$. Resuelva este problema: **a.** con el método de los multiplicadores de Lagrange; **b.** con ideas de geometría analítica.

30. En la elipse $4x^2 + 9y^2 = 36$, hallar el punto más cercano y el punto más alejado a la recta $x + y = 6$. (La distancia entre el punto (x, y) y la recta dada se toma perpendicularmente. Ver ejercicio 35 de la sección 3 del capítulo 1).

31. En la hipérbola $x^2 - y^2 = 1$, hallar los puntos más cercanos a la recta $y = 3x$.

32. En el elipsoide $4x^2 + y^2 + 4z^2 - 16x - 6y - 8z + 25 = 0$, hallar el punto más cercano y el punto más alejado del plano $2x + 2y + z = 0$. (Ver ejemplo 4).

33. En el paraboloide $z = (x - 2)^2 + 0.25(y - 3)^2 + 5$, hallar el punto más cercano al plano $x + y + z = 0$.

34. En el hiperboloide de una hoja $x^2 + \frac{1}{2}y^2 - \frac{1}{9}z^2 + 1 = 0$, hallar los puntos más cercanos al plano $x + y + 20(z - 1) = 0$.

35. Considere el elipsoide $6x^2 + 3y^2 + 2z^2 - 12x - 12y - 8z + 20 = 0$. Compruebe que su plano tangente en el punto $(1, 3, 0.5\sqrt{6} + 2)$ está dado por $\sqrt{6}z + 3y - 12 - 2\sqrt{6} = 0$. Sea $f(x, y, z)$ la función que da la distancia del punto $(x, y, z) \in \mathbb{R}^3$ al plano $\sqrt{6}z + 3y - 12 - 2\sqrt{6} = 0$. Determine los extremos de la función $f(x, y, z)$ con la restricción de que (x, y, z) es un punto del elipsoide dado.

36. Determine los semiejes de la elipse que se obtiene al intersectar el cilindro $x^2 + y^2 = 1$ con el plano $x + y + z = 0$, determinando los extremos condicionados de la función $f(x, y, z) = x^2 + y^2 + z^2$ sujeta a las dos restricciones $x^2 + y^2 - 1 = 0$, y $x + y + z = 0$.

37. Hallar el valor máximo y mínimo del producto de tres números reales x, y, z, si la suma de éstos debe ser cero y la suma de sus cuadrados debe ser uno.

(Distancia de un punto a una recta en $\mathbb{R}^3$). Sea $\mathbf{p} = (x_0, y_0, z_0)$ un punto dado en $\mathbb{R}^3$, y sea $\frac{x-x_1}{a} = \frac{y-y_1}{b} = \frac{z-z_1}{c}$ una recta L dada que pasa por (x_1, y_1, z_1) y tiene a (a, b, c) por vector paralelo. Se define la distancia de $\mathbf{p}$ a L, denotada por $d(\mathbf{p}, L)$, como la menor distancia que existe entre los puntos de la recta y el punto $\mathbf{p}$. Entonces, para calcular $d(\mathbf{p}, L)$ podemos considerar la función $f(x, y, z) = (x-x_0)^2+(y-y_0)^2+(z-z_0)^2 =$ (cuadrado de la) distancia del punto (x, y, z) al punto $\mathbf{p}$, con la restricción de que (x, y, z) se encuentre sobre L. Constate que esta última condición puede ser manejada como el par de restricciones: 1) $bx - ay + ay_1 - bx_1 = 0$; 2) $cx - az + az_1 - cx_1 = 0$. Use estas ideas para calcular la distancia de los puntos dados a las rectas dadas en los ejercicios 38–42.

38. $\mathbf{p} = (1, 0, 1), \dfrac{x-3}{2} = \dfrac{y-4}{3} = z$

39. $\mathbf{p} = (-2, 3, 2), \dfrac{x-1}{2} = \dfrac{y+1}{-2} = \dfrac{z+1}{2}$

40. $\mathbf{p} = (0, 0, 0), \dfrac{x-3}{5} = \dfrac{y+4}{-5} = \dfrac{z-3}{2}$

41. $\mathbf{p} = (0, 0, 1)$, L es la recta $\{(x, y, z)|x = y, z = 0\}$

42. $\mathbf{p} = (-1, -1, 3)$, L es la recta $\{(x, y, z)|x = y = z\}$

(Distancia entre dos rectas en $\mathbb{R}^3$). (Ver ejemplo 9 de la sección 6, capítulo 1). Sean L_1 y L_2 dos rectas en $\mathbb{R}^3$. Se define la distancia entre L_1 y L_2, denotada por $d(L_1, L_2)$, como la menor distancia entre los puntos (x, y, z) de L_1 y los puntos (x, y, z) de L_2. Si L_1 es la recta $\frac{x-x_1}{a} = \frac{y-y_1}{b} = \frac{z-z_1}{c}$ y L_2 la recta $\frac{x-x_2}{d} = \frac{y-y_2}{e} = \frac{z-z_2}{f}$, la distancia $d(L_1, L_2)$ se puede obtener como (la raíz cuadrada de) el mínimo

de la función $h(x, y, z, u, v, w) = (x-u)^2 + (y-v)^2 + (z-w)^2$, con las restricciones de que (x, y, z) esté en L_1 y (u, v, w) esté en L_2. Verifique entonces que $d(L_1, L_2)$ sea el mínimo condicionado de la función de 6 variables $h(x, y, z, u, v, w)$ sujeta a las 4 restricciones 1) $bx - ay + ay_1 - bx_1 = 0$; 2) $cx - az + az_1 - cx_1 = 0$; 3) $eu - dv + dy_2$ - $ex_2 = 0$; 4) $fu - dw + dz_2 - fx_2 = 0$. Use estas ideas para calcular la distancia entre las rectas L_1 y L_2 dadas en los ejercicios 43–46.

43. $\frac{x+4}{2} = \frac{y-4}{-1} = \frac{z+1}{-2}, \frac{x+5}{4} = \frac{y-5}{-3} = \frac{z-5}{-5}$

44. $x = y - 10 = z + 5, \frac{x+1}{2} = \frac{y+5}{3} = \frac{z-2}{4}$

45. $\frac{x+6}{1} = \frac{y+5}{2} = \frac{z+4}{3}, \frac{x+3}{4} = \frac{y+2}{5} = \frac{z+1}{6}$

46. $x = y = z, \frac{x+1}{2} = \frac{y-2}{-3} = \frac{z-1}{-4}$

47. Demostrar que el paralelepípedo de mayor volumen que se puede inscribir en una esfera es un cubo.

48. Hallar el paralelepípedo de mayor volumen que se puede inscribir en el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

49. Hallar el paralelepípedo de volumen V que tenga la menor superficie lateral posible.

50. Hallar el paralelepípedo de superficie lateral S que tenga el mayor volumen posible.

51. Hallar el paralelepípedo, cuya longitud de su diagonal d es dada, que tenga el mayor volumen posible.

52. Hallar el paralelepípedo cuya suma de longitudes de sus aristas sea L, que tenga el mayor volumen posible.

53. Considere la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Determine el triángulo de menor área que se puede formar (en el primer cuadrante) con la parte positiva de los ejes coordenados y una recta tangente a la elipse.

54. Considere el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Demuestre que el volumen del tetraedro formado por el plano tangente a este elipsoide en un punto de él (x, y, z) del primer octante y los planos coordenados, está dado por $V(x, y, z) = \frac{1}{6}\frac{a^2b^2c^2}{xyz}$. Minimize la función $V(x, y, z)$ sujeta a la restricción de que el punto (x, y, z) esté sobre el elipsoide, para concluir que el mínimo volumen que el tetredro mencionado puede tener es $V_{\text{mín}} = \frac{\sqrt{3}}{2}abc$. ¿En qué punto del elipsoide se debe trazar este plano tangente?

(*) **55.** Considere el elipsoide en $\mathbb{R}^n$ dado por la ecuación

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \cdots + \frac{x_n^2}{a_n^2} = 1$$

Básese en el hecho que el volumen de un $(n+1)$-edro en $\mathbb{R}^n$ formado por los hiperplanos coordenados y un plano que corta a los ejes coordenados en los puntos $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ está dado por $\frac{1}{n!}\bar{x}_1\bar{x}_2\ldots\bar{x}_n$ (ver ejemplo 2 de la sección 9 del capítulo 6), demuestre que el volumen del

$(n+1)$-edro formado por los hiperplanos coordenados ($x_i = 0, i = 1, 2, \ldots, n$) y un hiperplano tangente al elipsoide está dado por

$$V = V(x_1, x_2, \ldots, x_n) = \frac{1}{n!}\frac{a_1^2 a_2^2 \ldots a_n^2}{x_1 x_2 \ldots x_n}$$

Demuestre que el valor mínimo de esta función, con la restricción de que el punto $(x_1, x_2, \ldots, x_n)$ se encuentre en el elipsoide, es

$$V_{\text{mín}} = \frac{1}{n!}n^{n/2}a_1 a_2 \ldots a_n$$

NOTA: *Los siguientes ejercicios de determinación de extremos condicionados conducen a establecer sistemas no lineales (del tipo $\frac{\partial F}{\partial x_i} = 0$, donde F es la función de Lagrange) que no son fáciles de resolver "a mano", para los cuales se deben emplear técnicas numéricas para encontrar su solución, como el método de Newton estudiado en la sección 7 del capítulo 3.*

En los ejercicios 56–62 se dan ecuaciones de curvas cerradas en el plano. Se trata de determinar los puntos de las curvas que están más cercanos y más alejados del origen de las coordenadas. Se trata entonces de obtener los extremos de la función $f(x, y) = x^2 + y^2$ = cuadrado de la distancia del punto (x, y) al origen, con la restricción de que el punto (x, y) se encuentre en la curva dada. (Para dar el punto de inicio en el método de Newton, puede ayudar ver las gráficas de estas curvas).

56. $g(x, y) = x^4 + y^2 + xy - 4 = 0$

57. $g(x, y) = \ln(x^2 + 2x + 2) + \ln(y^2 + 2y + 2) - 1 = 0$

58. $g(x, y) = \ln(x^2 + 2x + 2) + y^2 - 6y + 8 = 0$

59. $g(x, y) = \ln(x^2 - 4x + 5) + y^2 - 6y + 8 = 0$

60. $g(x, y) = \exp(x^2 - y^2) + x^2 - 4x + y^2 - 6y + 10 = 0$

61. $g(x, y) = x^4 + y^4 + 2x^2 + 4y^2 + xy - 5 = 0$

62. $g(x, y) = (x^2 + y^2)^2 - 3x^2 + 4y^2 - 2xy - 2 = 0$

63. En el elipsoide $7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 18 = 0$ hallar el punto más cercano y más alejado del origen de coordenadas.

64. Hallar las distancias máxima y mínima entre los puntos de intersección del cilindro $x^2 + y^2 = 1$ con el plano $x + y + z = 0$, y los puntos de la recta $x = 2 + t$, $y = 3 - t$, $z = 6 + 5t$, $t \in \mathbb{R}$.

65. Hallar las distancias máxima y mínima entre los puntos de intersección del paraboloide $z = x^2 + 2y^2 - 1$ con el plano $x - y + z = 0$, y los puntos del plano $2x + 3y - 4z - 10 = 0$.

66. Hallar las distancias máxima y mínima entre los puntos del elipsoide $x^2 + 2y^2 + z^2 = 1$, y la recta $x = 3 - t$, $y = 4 + t$, $z = 6 + t$, $t \in \mathbb{R}$.

67. Hallar los puntos de la curva $x^2 + 4y^2 - 4 = 0$ que se encuentre más cercano y más alejado de los puntos de la curva $x^2 + y^2 + 4x + 2y - 20 = 0$.

Los ejercicios 68–74 tratan acerca de extremos absolutos de funciones continuas en regiones compactas, tema estudiado en el apéndice de esta sección.

68. Hallar los extremos absolutos de la función $f(x, y) = \frac{1}{3}x^3 - \frac{3}{2}x^2 + 2x + y^2 - 2y + 1$, en la región $K = \{(x, y)|x \geq 0, y \geq 0, x + y \leq 1\}$.

69. Hallar los extremos absolutos de la función $f(x, y) = x + y$, en la región $K = \{(x, y)| -1 \leq x \leq 1, -1 \leq y \leq 1\}$.

70. Determinar los extremos absolutos de la función $f(x, y) = 3x^2 + 5y^2$, en la región $K = \{(x, y)|0 \leq x \leq 1, 0 \leq y \leq 1\}$.

71. Determinar los extremos absolutos de la función $f(x, y) = x^2 + x + y^2 + y$, en la región $K = \{(x, y)|x^2 + y^2 \leq 1\}$.

72. Determinar los extremos absolutos de la función $f(x, y) = \cos x + \cos y$ en la región $K = \{(x, y)| -\pi \leq x \leq \pi, -\pi \leq y \leq \pi\}$.

73. Hallar los extremos absolutos de la función $f(x, y) = \operatorname{sen} x + \cos y$ en la región $K = \{(x, y)| -\pi \leq x \leq \pi, -\pi \leq y \leq \pi\}$.

74. Hallar los extremos absolutos de la función $f(x, y, z) = (x - 1)^2 + (y - 1)^2 + (z - 1)^2$, en la región $K = \{(x, y, z)|x^2 + y^2 + z^2 \leq 12\}$.

(*)4.6 Extremos condicionados (II). Condiciones suficientes

El objetivo de esta sección es, por una parte señalar por medio de ejemplos algunos de los detalles finos que tiene el método de los multiplicadores de Lagrange expuesto en la sección anterior, y , por otra parte, asomarnos un poco en la dirección de la obtención de condiciones suficientes para la existencia de extremos condicionados. El material aquí presentado es opcional (para un primer curso de Cálculo Vectorial) y se puede omitir sin alterar en absoluto la secuencia de los temas; lo estudiado en la sección anterior es la parte estándar del tema de extremos condicionados que debe ser incluido en cualquier curso de esta materia.

Consideremos en primer lugar la siguiente situación: si quisiéramos obtener los extremos de la función $z = f(x, y)$ sujeta a la restricción $g(x, y) = 0$, y de alguna manera nos la pudiéramos ingeniar para despejar (de esta última expresión), a y como función de x, dejando así establecida la función $y = \varphi(x)$, la que se sustituye entonces en f, finalmente logramos así convertir nuestro problema inicial en el de la investigación de extremos ("normales", de los del primer curso de cálculo) para la función $z = f(x, \varphi(x)) = F(x)$ que depende sólo de la variable x. Todo esto suena muy bien, pero podrían surgir algunos problemas. Veamos el siguiente ejemplo.

Ejemplo 1. Se quiere encontrar la distancia más corta entre los puntos de la parábola $y^2 = x - 1$ y el origen. Consideramos entonces la función $f(x, y) = x^2 + y^2$, que nos da (el cuadrado de) la distancia del punto $(x, y) \in \mathbb{R}^2$ al origen, e imponemos la restricción de que este punto pertenezca a la parábola $y^2 = x - 1$. Nuestra restricción es entonces $g(x, y) = y^2 - x + 1 = 0$. Una sencilla reflexión geométrica de la siguiente figura nos dice que la solución debe ser el punto $(1, 0)$
Sin embargo, si aplicamos nuestra técnica de convertir el problema en el de una función de una sola variable, vemos que fácilmente de $g(x, y) = y^2 - x + 1 = 0$ despejamos y en términos de x, poniendo $y^2 = x - 1$ y sustituimos en la función f, obteniendo la función $f(x) = x^2 + x - 1$, de la cual,

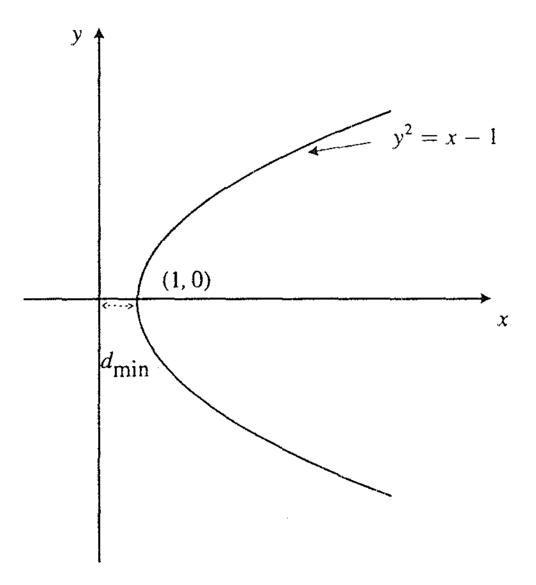


Figura 1. La parábola $y^2 = x - 1$ del ejemplo 1.

nos disponemos a localizar los extremos. De $F'(x) = 2x + 1 = 0$ obtenemos $x = -\frac{1}{2}$, y así nos encontramos con la desagradable sorpresa de que no existe punto alguno de la parábola $y^2 = x - 1$ que tenga esta abscisa (no existe y por tanto $y^2 = -\frac{1}{2} - 1 = -\frac{3}{2}$). La explicación de este conflicto es sencilla: en el punto $\mathbf{p} = (1, 0)$ en el que se encuentra la solución del problema no se puede ver a y como función de x (en una vecindad de él). Esto es lo que nos indica el Teorema de la función Implícita (Teorema 3.4.1). Obsérvese que $\frac{\partial g}{\partial y}(1, 0) = 0$, de modo que de $g(x, y) = 0$ no podemos poner $y = \varphi(x)$ para una vecindad de $(1, 0)$. Sin embargo, a la luz de este mismo teorema de funciones implícitas, como $\frac{\partial g}{\partial x}(1, 0) = -1 \neq 0$, sí podemos poner a x como función de y en una vecindad de $\mathbf{p}$. Dicha función es $x = y^2 + 1$. Si sustituimos ésta en la función f, obtenemos $F(y) = (y^2 + 1)^2 + y^2$, investigando los extremos de esta función tenemos: $F'(y) = 2(y^2 + 1)2y + 2y = 0$ de donde se obtiene $y = 0$, y como, $x = y^2 + 1$, se obtiene $x = 1$, llegando así a la solución del problema. La moraleja de este ejemplo es que el método para despejar una variable en términos de la otra de la restricción $g(x, y) = 0$ para luego sustituir en la función de la cual se quieren obtener los extremos, resulta riesgoso. Peor aún: no hay manera de saber a priori si tal despeje es lícito (a la luz del teorema de la función implícita) pues para eso tendríamos que saber de antemano la solución del problema. Ante esta desconsoladora situación, surge el método de los multiplicadores de Lagrange, que lo único que pide es que grad $g(\bar{x}, \bar{y}) \neq \mathbf{0}$ (el vector cero) en el punto $(\bar{x}, \bar{y})$ en el que se encuentra el extremo. Como en nuestro caso grad $g(x, y) = (2y, -1) \neq (0, 0)\ \forall (x, y) \in \mathbb{R}^2$, se puede proceder sin ningún temor con este método:

$$F(x, y, \lambda) = x^2 + y^2 + \lambda(y^2 - x + 1)$$

de donde

$$\frac{\partial F}{\partial x} = 2x - \lambda = 0, \quad \frac{\partial F}{\partial y} = 2y + 2\lambda y = 0, \quad \frac{\partial F}{\partial \lambda} = y^2 - x + 1 = 0$$

La única solución para este sistema es $x = 1$, $y = 0$, que es la solución de nuestro problema. ■

Al estudiar los extremos de la función $z = f(x, y)$ sujeta a la restricción $g(x, y) = 0$, la hipótesis de que grad $g(\bar{x}, \bar{y}) \neq (0, 0)$ en el punto $(\bar{x}, \bar{y})$ en donde se encuentra el extremo procurado es importante, como se ve en el ejemplo siguiente.

Ejemplo 2. Se desea obtener los extremos de la función $f(x, y) = x^2 + y^2$ sujeta a la restricción $g(x, y) = (x-1)^3 - y^2 = 0$. Apliquemos el método de los multiplicadores de Lagrange para resolver nuestro problema. La función de Lagrenge en este caso es

$$F(x, y, \lambda) = x^2 + y^2 + \lambda((x-1)^3 - y^2)$$

de donde

$$\frac{\partial F}{\partial x} = 2x + 3\lambda(x-1)^2 = 0, \qquad \frac{\partial F}{\partial y} = 2y - 2\lambda y = 0$$

$$\frac{\partial F}{\partial \lambda} = (x-1)^3 - y^2 = 0$$

Este sistema no tiene solución alguna. Sin embargo, si pensamos en términos geométricos, vemos que lo que se intenta es obtener el mínimo y/o el máximo de la función distancia de los puntos de la curva $g(x, y) = 0$ al origen. Esta curva es $y^2 = (x-1)^3$ cuya gráfica es la figura 2 de donde se ve claramente que en el punto $(1, 0)$ se debe tener un mínimo para la función $f(x, y) = x^2 + y^2$. El problema es que en ese punto se tiene

$$\text{grad}\, g(1, 0) = (3(x-1)^2, -2y)\Big|_{\substack{x=1\\y=0}} = (0, 0)$$

Así que en los alrededores del punto $(1, 0)$, la curva $g(x, y) = 0$ no tiene un comportamiento "decente", violándose entonces una hipótesis fundamental del método de los multiplicadores de Lagrange. ■

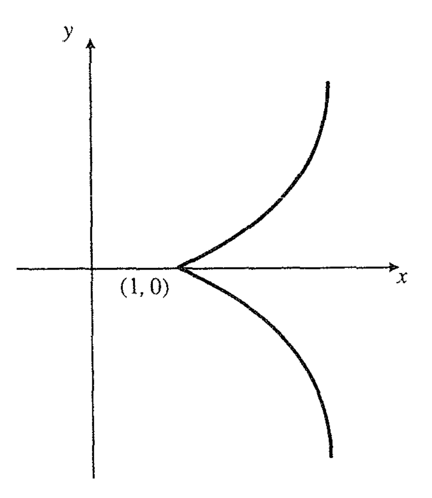


Figura 2. Gráfica de la curva $y^2 = (x-1)^3$.

Al considerar el problema de estudiar los extremos de la función $u = f(x, y, z)$ sujeta a dos restricciones $g_1(x, y, z) = 0$ y $g_2(x, y, z) = 0$, se pidió que en el punto **p**, donde se de el extremo, se tuvieran los vectores grad $g_1(\mathbf{p})$ y grad $g_2(\mathbf{p})$ linealmente independientes. Esta es en cierto sentido la misma conducción de que grad $g(\bar{x}, \bar{y}) \neq 0$ comentada previamente. En cada una de las situaciones en las que aparecen tales condiciones, ellas garantizan la existencia de los multiplicadores de Lagrange λ_i, que es la tesis fundamental en la que se basa el método.

En el teorema general estas condiciones están incluidas en la hipótesis de que si $\mathbf{x}_0 \in \mathbb{R}^n$ en el punto donde se de el extremo de la función de n variables $z = f(x_1, x_2, \ldots, x_n)$ sujeta a las m restricciones $g_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, m$, entonces debe cumplirse que

$$\det\left(\frac{\partial g_i}{\partial x_j}(\mathbf{x}_0)\right) \neq 0 \quad i, j = 1, 2, \ldots m$$

para alguna elección de m variables x_j tomadas del conjunto de las n variables $(x_1, x_2, \ldots, x_n)$ (se entiende que $m < n$).

En efecto, si $n = 2$, $m = 1$, estamos en el caso de una función de dos variables $z = f(x, y)$ sujeta a una restricción $g(x, y) = 0$. La hipótesis anterior pide que alguno de los determinantes

$$\det\left(\frac{\partial g}{\partial x}\right) = \frac{\partial g}{\partial x}, \quad \det\left(\frac{\partial g}{\partial y}\right) = \frac{\partial g}{\partial y}$$

sea no nulo en el punto $(\bar{x}, \bar{y})$ donde se de el extremo. Esto significa que el vector grad $g = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)$ no se anule en $(\bar{x}, \bar{y})$ —que no tenga ambas componentes igual a cero—, la cual es justamente la condición que conocíamos para este caso.

Si $n = 3$, $m = 2$, estamos en el caso de una función de tres variables $u = f(x, y, z)$ sujeta a dos restricciones $g_1(x, y, z) = 0$, $g_2(x, y, z) = 0$. La hipótesis general pide que alguno de los determinantes

$$\det\begin{bmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{bmatrix} = \frac{\partial(g_1, g_2)}{\partial(x, y)}$$

$$\det\begin{bmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial z} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial z} \end{bmatrix} = \frac{\partial(g_1, g_2)}{\partial(x, z)}$$

$$\det\begin{bmatrix} \frac{\partial g_1}{\partial y} & \frac{\partial g_1}{\partial z} \\ \frac{\partial g_2}{\partial y} & \frac{\partial g_2}{\partial z} \end{bmatrix} = \frac{\partial(g_1, g_2)}{\partial(y, z)}$$

(que son los jacobianos de g_1 y g_2 respecto a x y y, x y z y y y z respectivamente) sea no nulo en el punto **p** donde se da el extremo.

Por otra parte, sabemos que los dos vectores $\mathbf{v}_1$ = grad g_1 y $\mathbf{v}_2$ = grad g_2 en $\mathbb{R}^3$ son linealmente independientes si y sólo si su producto cruz $\mathbf{v}_1 \times \mathbf{v}_2$ es no nulo. Tal producto cruz es

$$\operatorname{grad} g_1 \times \operatorname{grad} g_2 = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} & \dfrac{\partial g_1}{\partial z} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} & \dfrac{\partial g_2}{\partial z} \end{bmatrix}$$

$$= \frac{\partial(g_1, g_2)}{\partial(y, z)}\mathbf{i} - \frac{\partial(g_1, g_2)}{\partial(x, z)}\mathbf{j} + \frac{\partial(g_1, g_2)}{\partial(x, y)}\mathbf{k}$$

el cual es no nulo si y sólo si alguno de sus componentes es no nulo. Por lo tanto, la hipótesis del teorema general es, en este caso, justamente la condición de la independencia lineal de los vectores grad g_1 y grad g_2 que ya conocemos.

Ejemplo 3. Consideremos el problema de encontrar los extremos de la función $u = x^2 + y^2 + z^2$ sujeta a las dos restricciones $g_1(x, y, z) = z - (y-1)^3 = 0$ y $g_2(x, y, z) = z = 0$. Tenemos que la función de Lagrange es en este caso

$$F(x, y, z, \lambda_1, \lambda_2) = x^2 + y^2 + z^2 + \lambda_1(z - (y-1)^3) + \lambda_2 z$$

de donde

$$\frac{\partial F}{\partial x} = 2x = 0 \qquad (1)$$

$$\frac{\partial F}{\partial y} = 2y - 3\lambda_1(y-1)^2 = 0 \qquad (2)$$

$$\frac{\partial F}{\partial z} = 2z + \lambda_1 + \lambda_2 = 0 \qquad (3)$$

$$\frac{\partial F}{\partial \lambda_1} = z - (y-1)^3 = 0 \qquad (4)$$

$$\frac{\partial F}{\partial \lambda_2} = z = 0 \qquad (5)$$

Este sistema es incompatible pues de (4) y (5) se sigue que el valor de y debería ser 1, que no satisface (2). Viendo el problema en forma geométrica, se trata de hallar el máximo y el mínimo (si existen) de la función distancia (al cuadrado) del punto (x, y, z) al origen, $f(x, y, z) = x^2 + y^2 + z^2$, con la restricción de que el punto (x, y, z) se encuentra en la intersección de las dos superficies $g_1(x, y, z) = z - (y-1)^3 = 0$ y $g_2(x, y, z) = z = 0$. Tal intersección es la recta $x = t$, $y = 1$, $z = 0$, $t \in \mathbb{R}$. Geométricamente es claro que $(0, 1, 0)$ es el punto de esta recta que se encuentra más cerca del origen. Esta debe ser la solución de nuestro problema que, en este caso, no se detecta por el método de los multiplicadores de Lagrange.

El conflicto en este problema está en que en el punto $\mathbf{p} = (0, 1, 0)$, que es donde se alcanza el mínimo procurado, se tiene

$$\operatorname{grad} g_1(0, 1, 0) = (0, -3(y-1)^1, 1)\Big|_{\substack{x=0\\y=1\\z=0}} = (0, 0, 1)$$

$$\operatorname{grad} g_2(0, 1, 0) = (0, 0, 1)$$

de modo que los dos vectores grad g_1 y grad g_2 son linealmente dependientes en ese punto ■

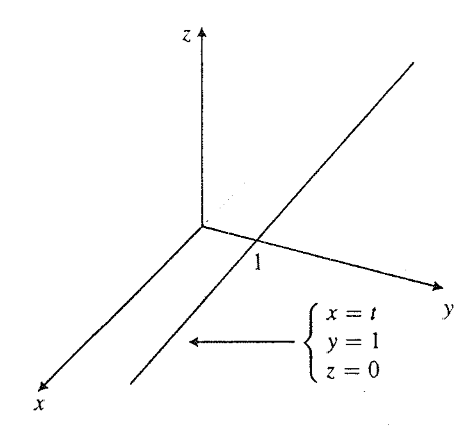


Figura 3. Gráfica de la recta $x = t$, $y = 1$, $z = 0$, $t \in \mathbb{R}$.

Quisiéramos ahora pasar a discutir un poco sobre las condiciones suficientes para la existencia de los extremos condicionados. Este es un tema suficientemente complicado como para superar las pretenciones de cualquier libro estándar de Cálculo Avanzado. No será nuestra intención por lo tanto discutirlo con profundidad. Sin embargo, creemos que siempre es intelectualmente reconfortante, darse cuenta al menos del tipo de dificultades al respecto y hacer algunas cuentas, en un caso particular, para "sentir" tales dificultades.

Comencemos por considerar un ejemplo concreto. Tomemos la función $f(x, y) = y - x^4$ sujeta a la restricción $g(x, y) = y - x^3 = 0$. La función de Lagrange es en este caso

$$F(x, y, \lambda) = y - x^4 + \lambda(y - x^3)$$

de donde

$$\frac{\partial F}{\partial x} = -4x^3 - 3\lambda x^2 = 0, \quad \frac{\partial F}{\partial y} = 1 + \lambda = 0, \quad \frac{\partial F}{\partial \lambda} = y - x^3 = 0$$

El origen es un punto cuyas coordenadas satisfacen este sistema. Sin embargo, podemos ver que en tal punto no puede existir un extremo de la función f. En efecto, consideremos la curva de nivel $f(x, y) = f(0, 0)$, o sea $y = x^4$, la cual separa los puntos del plano en los que $f(x, y) > f(0, 0)$ de los puntos en los que $f(x, y) < f(0, 0)$. Veamos qué ocurre cerca del origen con esta curva y la curva $y = x^3$ que marca la restricción del problema

Se ve que cualquier bola con centro en el origen contiene puntos de la curva $y = x^3$ para los que $f(x, y) > f(0, 0)$ y puntos de esa curva para los que $f(x, y) < f(0, 0)$

Así pues, el origen es un punto que cumple con las condiciones necesarias para la existencia de un extremo condicionado de la función f, en el que se muestra que tales condiciones no son suficientes.

Tomemos ahora en general una función $z = f(x, y)$ de clase $\mathscr{C}^1$ (como lo pide el teorema de los multiplicadores de Lagrange) y una función $z = g(x, y)$ también de clase $\mathscr{C}^1$, e investiguemos los extremos de la función f sujeta a la restricción $g(x, y) = 0$.

Ya sabemos qué hacer para obtener un punto $\mathbf{p} = (x_0, y_0)$ candidato a extremo condicionado. Haremos ahora algunas operaciones que nos permitan obtener en este caso condiciones suficientes para que en el punto $\mathbf{p}$ exista de hecho como extremo. Sabemos que en el punto $\mathbf{p}$ el vector grad g

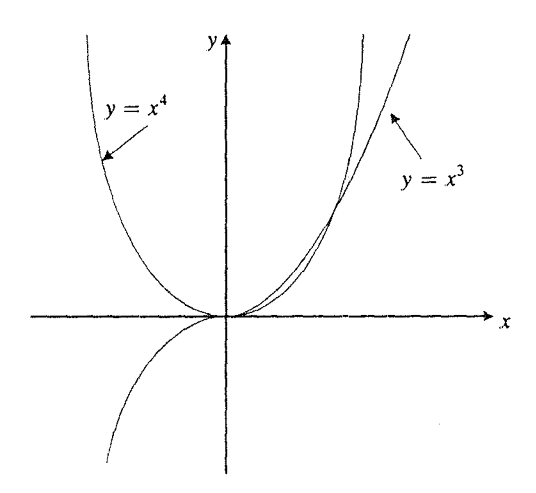


Figura 4. Gráfica de $y = x^4$ y $y = x^3$.

es no nulo. Supongamos, por ejemplo, que

$$\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$$

El teorema de la función implícita nos dice que, en tal caso, podemos ver la curva $g(x, y) = 0$, en los alrededores del punto **p**, como la gráfica de una función de clase $\mathscr{C}^1$, $y = \varphi(x)$, cuya derivada se puede calcular como

$$y' = \varphi'(x) = -\frac{\dfrac{\partial g}{\partial x}}{\dfrac{\partial g}{\partial y}}$$

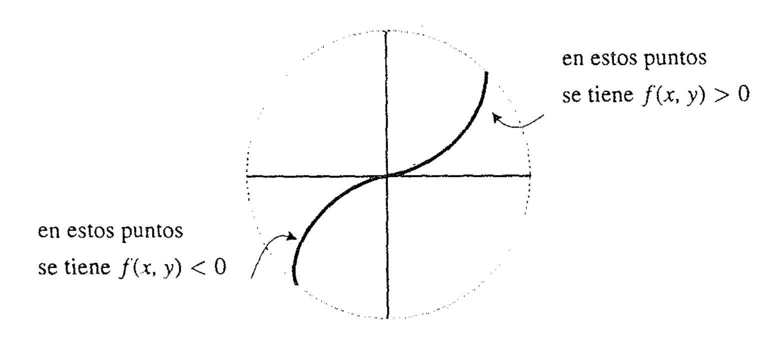


Figura 5. Gráfica de $y = x^3$ en una bola con centro en el origen.

para los puntos que están cerca de **p**. En tal caso, podemos considerar la función $z = f(x, \varphi(x))$ que, sabemos, tiene un punto crítico para $x = x_0$. En efecto, se tiene

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\varphi'(x) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\left(-\frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}}\right)$$
$$= \frac{1}{\frac{\partial g}{\partial y}}\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right)$$

de modo que $\frac{dz}{dx}(x_0) = 0$ si y solo si

$$\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} = 0$$

en el punto (x_0, y_0), donde $y_0 = \varphi(x_0)$. Esta relación es equivalente a la existencia de $\lambda \in \mathbb{R}$ tal que

$$\operatorname{grad} f + \lambda \operatorname{grad} g = 0$$

lo cual está expresado por el teorema de los multiplicadores de Lagrange. Más aún, la función $z = f(x, \varphi(x))$ tendrá un extremo local en $x = x_0$ si y sólo si la función $z = f(x, y)$ tiene tal extremo (condicionado, con la restricción $g(x, y) = 0$) en el punto $\mathbf{p} = (x_0, y_0)$. Invitamos al lector para que tome de la definición de extremos locales para $z = f(x, \varphi(x))$ y de extremos condicionados para $z = f(x, y)$ con la restricción $g(x, y) = 0$, para que se convenza plenamente de esta afirmación. Lo que pretendemos entonces es usar el criterio de la segunda derivada para identificar extremos locales de la función $z = f(x, \varphi(x))$ y dejarlo en términos de las funciones f y g que conocemos en nuestro problema. Todo consiste entonces en realizar algunas operaciones haciendo uso intensivo de la regla de la cadena.

Tenemos ya calculada $\frac{dz}{dx}$. Esta es

$$\frac{dz}{dx} = \frac{\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}}$$

Cada una de las derivadas que aparecen en el miembro derecho de esta expresión son funciones de las variables x y y. Denotémoslas en general como $\Phi(x, y)$. Suponemos además que $y = \varphi(x)$. Tendremos que derivar $\Phi(x, y)$ respecto de x. Entonces, el esquema general será entonces según la regla de la cadena

$$\frac{\partial}{\partial x}\Phi(x, y) = \frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y}\varphi'(x) = \frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y}\left(-\frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}}\right)$$

Procedamos entonces a derivar nuevamente $\frac{dz}{dx}$ respecto de x. Tenemos

$$\frac{d^2z}{dx^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}}\right)$$

$$= \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left\{\frac{\partial g}{\partial y}\left[\frac{\partial f}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y}\right) + \frac{\partial g}{\partial y}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) - \frac{\partial f}{\partial y}\frac{\partial}{\partial x}\left(\frac{\partial g}{\partial x}\right)\right.\right.$$

$$\left.\left. - \frac{\partial g}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right] - \left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right)\frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y}\right)\right\}$$

Aplicando nuestro esquema general, comentado previamente, para calcular las derivadas indicadas obtenemos

$$\frac{d^2z}{dx^2} = \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left[\frac{\partial^2 g}{\partial x\partial y}\frac{\partial g}{\partial y}\frac{\partial f}{\partial x} - \frac{\partial^2 g}{\partial y^2}\frac{\partial g}{\partial x}\frac{\partial f}{\partial x} + \left(\frac{\partial g}{\partial y}\right)^2\frac{\partial^2 f}{\partial x^2}\right.$$

$$\left. - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 f}{\partial x\partial y} - \frac{\partial^2 g}{\partial x^2}\frac{\partial g}{\partial y}\frac{\partial f}{\partial y} + \frac{\partial^2 g}{\partial x\partial y}\frac{\partial g}{\partial x}\frac{\partial f}{\partial y} + \left(\frac{\partial g}{\partial x}\right)^2\frac{\partial^2 f}{\partial y^2}\right]$$

Todas las derivadas se tienen calculadas en (x_0, y_0). Aceptamos también que f y g son funciones de clase $\mathscr{C}^2$, de modo que sus parciales cruzadas son iguales. Esta es la expresión que buscábamos para $\frac{d^2z}{dx^2}$ en términos de las derivadas de las funciones conocidas f y g. Si esta expresión es positiva en (x_0, y_0) podemos concluir la existencia de un mínimo condicionado para la función f, y si es negativa, de un máximo condicionado.

Es posible, con un poco de trabajo adicional, hacer que el aspecto de esta expresión sea menos terrible, si en lugar de las funciones f y g, trabajamos con la función f y con la función de Lagrange $F = f + \lambda g$. En el punto $\mathbf{p} = (x_0, y_0)$ se tiene, para λ = valor del multiplicador de Lagrange correspondiente a $\mathbf{p}$

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + \lambda\frac{\partial g}{\partial x} = 0, \quad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} + \lambda\frac{\partial g}{\partial y} = 0,$$

$$\frac{\partial^2 F}{\partial x^2} = \frac{\partial^2 f}{\partial x^2} + \lambda\frac{\partial^2 g}{\partial x^2}, \quad \frac{\partial^2 F}{\partial x\partial y} = \frac{\partial^2 f}{\partial x\partial y} + \lambda\frac{\partial^2 g}{\partial x\partial y}, \quad \frac{\partial^2 F}{\partial y^2} = \frac{\partial^2 f}{\partial y^2} + \lambda\frac{\partial^2 g}{\partial y^2}$$

de donde se tienen las expresiones

$$\frac{\partial f}{\partial x} = -\lambda\frac{\partial g}{\partial x}, \quad \frac{\partial f}{\partial y} = -\lambda\frac{\partial g}{\partial y},$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 F}{\partial x^2} - \lambda\frac{\partial^2 g}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 F}{\partial x\partial y} - \lambda\frac{\partial^2 g}{\partial x\partial y}, \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 F}{\partial y^2} - \lambda\frac{\partial^2 g}{\partial y^2}$$

Sustituyendo éstas en la expresión para $\frac{d^2z}{dx^2}$ nos queda

$$\begin{aligned}\frac{d^2z}{dx^2} &= \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left[\frac{\partial^2 g}{\partial x\partial y}\frac{\partial g}{\partial y}\left(-\lambda\frac{\partial g}{\partial x}\right) - \frac{\partial^2 g}{\partial y^2}\frac{\partial g}{\partial x}\left(-\lambda\frac{\partial g}{\partial x}\right) + \left(\frac{\partial g}{\partial y}\right)^2\left(\frac{\partial^2 F}{\partial x^2} - \lambda\frac{\partial^2 g}{\partial x^2}\right)\right.\\ &\quad - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\left(\frac{\partial^2 F}{\partial x\partial y} - \lambda\frac{\partial^2 g}{\partial x\partial y}\right) - \frac{\partial^2 g}{\partial x^2}\frac{\partial g}{\partial y}\left(-\lambda\frac{\partial g}{\partial y}\right)\\ &\quad \left. + \frac{\partial^2 g}{\partial x\partial y}\frac{\partial g}{\partial x}\left(-\lambda\frac{\partial g}{\partial y}\right) + \left(\frac{\partial g}{\partial x}\right)^2\left(\frac{\partial^2 F}{\partial y^2} - \lambda\frac{\partial^2 g}{\partial y^2}\right)\right]\\ &= \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left[\left(\frac{\partial g}{\partial x}\right)^2\frac{\partial^2 F}{\partial y^2} - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 F}{\partial x\partial y} + \left(\frac{\partial g}{\partial y}\right)^2\frac{\partial^2 F}{\partial x^2}\right]\end{aligned}$$

Observamos que la expresión dentro del corchete es el negativo del desarrollo del determinante

$$H_F = \det\begin{bmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x\partial y} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial y^2}\end{bmatrix}$$

que llamamos *hessiano limitado* de la función $F = f + \lambda g$ en el punto $\mathbf{p}$. Así pues

$$\frac{\partial^2 z}{\partial x^2}(x_0) = -\frac{H_F(x_0, y_0)}{\left(\frac{\partial g}{\partial y}(x_0, y_0)\right)^2}$$

Como el denominador de esta expresión es siempre positivo, el signo de $\frac{d^2z}{dx^2}(x_0)$ queda determinado por el signo de $H_F(x_0, y_0)$, teniéndose que:

(*) Si $H_F(x_0, y_0) > 0$ $\left(\text{en cuyo caso } \frac{d^2z}{dx^2}(x_0) < 0\right)$, entonces la función f tiene un máximo condicionado en (x_0, y_0);

(*) Si $H_F(x_0, y_0) < 0$ $\left(\text{en cuyo caso } \frac{d^2z}{dx^2}(x_0) > 0\right.$, entonces la función f tiene un mínimo condicionado en $\left.(x_0, y_0)\right)$;

(*) Si $H_F(x_0, y_0) = 0$ no podemos determinar la naturaleza del punto (x_0, y_0).

Veamos un ejemplo.

Ejemplo 4. Se quieren estudiar los extremos de la función $f(x, y) = \cos^2 x + \cos^2 y$ sujeta a la restricción $g(x, y) = x - y - \frac{\pi}{4} = 0$. La función de Lagrange es

$$F(x, y, \lambda) = \cos^2 x + \cos^2 y + \lambda(x - y - \frac{\pi}{4})$$

de donde

$$\frac{\partial F}{\partial x} = -\operatorname{sen} 2x + \lambda = 0 \tag{1}$$

$$\frac{\partial F}{\partial y} = -\operatorname{sen} 2y - \lambda = 0 \tag{2}$$

$$\frac{\partial F}{\partial \lambda} = x - y - \frac{\pi}{4} = 0 \tag{3}$$

De (1) y (2) se tiene que $\operatorname{sen} 2x + \operatorname{sen} 2y = 0$. Incorporando la ecuación (3) obtenemos $\tan 2x = 1$, de donde $x = \frac{\pi}{8} + k\frac{\pi}{2}$, $k \in \mathbb{Z}$, y por lo tanto, de (3), $y = -\frac{\pi}{8} + k\frac{\pi}{2}$, $k \in \mathbb{Z}$, se tienen así una infinidad de posibles extremos, a saber

$$\begin{aligned} x &= \frac{\pi}{8} + k\frac{\pi}{2} \\ y &= -\frac{\pi}{8} + k\frac{\pi}{2} \end{aligned} \qquad k \in \mathbb{Z}$$

Obtengamos los elementos del hessiano limitado para la función F. Estos son

$$\frac{\partial g}{\partial x} = 1, \quad \frac{\partial g}{\partial y} = -1$$

$$\frac{\partial^2 F}{\partial x^2} = -2\cos 2x, \quad \frac{\partial^2 F}{\partial x \partial y} = 0, \quad \frac{\partial^2 F}{\partial y^2} = -2\cos 2y$$

Evaluando en los puntos $\mathbf{p}_k = \left(\frac{\pi}{8} + k\frac{\pi}{2}, -\frac{\pi}{8} + k\frac{\pi}{2}\right)$ obtenemos

$$\frac{\partial^2 F}{\partial x^2}(\mathbf{p}_k) = -2\cos\left(\frac{\pi}{4} + k\pi\right) = -2(-1)^k\frac{\sqrt{2}}{2} = (-1)^{k+1}\sqrt{2}$$

$$\frac{\partial^2 F}{\partial y^2}(\mathbf{p}_k) = -2\cos\left(-\frac{\pi}{4} + k\pi\right) = -2(-1)^k\frac{\sqrt{2}}{2} = (-1)^{k+1}\sqrt{2}$$

de modo que el hessiano limitado H_F es

$$H_F = \det\begin{bmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial y^2} \end{bmatrix} = \det\begin{bmatrix} 0 & -1 & 1 \\ -1 & (-1)^{k+1}\sqrt{2} & 0 \\ 1 & 0 & (-1)^{k+1}\sqrt{2} \end{bmatrix} = (-1)^k 2\sqrt{2}$$

Concluimos entonces que si k es par ($H_F > 0$), la función f tendrá máximos condicionados en los puntos $\mathbf{p}_k$, que valen

$$f(p_k) = \cos^2\left(\frac{\pi}{8} + k\frac{\pi}{2}\right) + \cos^2\left(-\frac{\pi}{8} + k\frac{\pi}{2}\right) = 1 + \frac{\sqrt{2}}{2}$$

pero si k es impar ($H_F < 0$), la función f tendrá mínimos condicionados en los puntos $\mathbf{p}_k$ que valen

$$f(\mathbf{p}_k) = 1 - \frac{\sqrt{2}}{2}$$

■

Tomemos ahora el caso de una función $f(x, y, z)$ de tres variables x, y, z, sujeta a la restricción $g(x, y, z) = 0$. Sea (x_0, y_0, z_0) un punto candidato a extremo condicionado de f, obtenido según el método de los multiplicadores de Lagrange. Sabemos que grad $g(\mathbf{p}) \neq 0$. Sin perder generalidad podemos suponer que $\frac{\partial g}{\partial y}(\mathbf{p}) \neq 0$. Esto significa entonces que (por el teorema de la función implícita) en una vecindad del punto $\mathbf{p}$ podemos ver la superficie $g(x, y, z) = 0$ como la gráfica de una función $y = \varphi(x, z)$. Sustituyendo ésta en la función f, obtenemos la función de dos variables

$$u = f(x, \varphi(x, z), z)$$

Es fácil ver que esta función tiene un punto crítico en (x_0, z_0), pues este hecho equivale a la existencia del multiplicador de Lagrange λ de modo que grad $f(\mathbf{p}) + \lambda$ grad $f(\mathbf{p}) = \mathbf{0}$ $(\in \mathbb{R}^3)$. Más aún, la función $u = f(x, \varphi(x, z), z)$ tendrá un extremo local en (x_0, z_0) si y solamente si la función $u = f(x, y, z)$ tiene tal tipo de extremo (condicionado con la restricción $g(x, y) = 0$) en el punto $\mathbf{p} = (x_0, y_0, z_0)$. El problema original se convierte en el de la determinación de extremos locales de la función $u = f(x, \varphi(x, z), z)$. Siendo ésta una función de dos variables, podemos hacer uso del Teorema 3.4.1 que nos da condiciones suficientes para la existencia de extremos locales para este tipo de funciones. En nuestro caso tal teorema nos daría condiciones suficientes para la existencia de extremos condicionados de la función f en términos de las derivadas

$$A = \frac{\partial^2 u}{\partial x^2}(x_0, z_0), \quad B = \frac{\partial^2 u}{\partial x \partial z}(x_0, z_0), \quad C = \frac{\partial^2 u}{\partial z^2}(x_0, z_0)$$

Tal como hicimos en el caso anterior, debemos de tener estas derivadas en términos de las (derivadas de) las funciones g y f (ó g y $F = f + \lambda g$), para que el criterio nos sea de utilidad.

Las cuentas que se hacen para obtener A, B y C son muy parecidas a las realizadas en el caso anterior. Se trata simplemente de aplicar de manera correcta la regla de la cadena y los resultados del teorema 3.4.2 para la derivación de las funciones implícitas (al momento de obtener $\frac{\partial y}{\partial x}$ y $\frac{\partial y}{\partial z}$ en términos de las derivadas parciales de la función g). Damos solamente los resultados, en términos de la función g y de la función de Lagrange $F = f + \lambda g$. Se supone también que tanto f como g son funciones de clase $\mathscr{C}^2$, para simplificar las expresiones resultantes tomando iguales las parciales mixtas correspondientes. Se tiene

$$A = \frac{\partial^2 u}{\partial x^2} = \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left[\left(\frac{\partial g}{\partial x}\right)^2 \frac{\partial^2 F}{\partial y^2} - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 F}{\partial x \partial y} + \left(\frac{\partial g}{\partial y}\right)^2 \frac{\partial^2 F}{\partial x^2}\right]$$

$$C = \frac{\partial^2 u}{\partial z^2} = \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left[\left(\frac{\partial g}{\partial x}\right)^2 \frac{\partial^2 F}{\partial z^2} - 2\frac{\partial g}{\partial y}\frac{\partial g}{\partial z}\frac{\partial^2 F}{\partial y \partial z} + \left(\frac{\partial g}{\partial z}\right)^2 \frac{\partial^2 F}{\partial y^2}\right]$$

$$B = \frac{\partial^2 f}{\partial x \partial y} = \frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\left[\left(\frac{\partial g}{\partial y}\right)^2 \frac{\partial^2 F}{\partial x \partial z} - \frac{\partial g}{\partial y}\frac{\partial g}{\partial z}\frac{\partial^2 F}{\partial x \partial y} - \frac{\partial g}{\partial y}\frac{\partial g}{\partial x}\frac{\partial^2 f}{\partial z \partial y} + \frac{\partial g}{\partial x}\frac{\partial f}{\partial z}\frac{\partial^2 F}{\partial y^2}\right]$$

Se supone que todas las derivadas están evaluadas en $\mathbf{p}$. Al calcular la expresión $B^2 - AC$ (la cual debe ser negativa para que pueda existir extremo) nos encontramos con una agradable sorpresa. Este

es justamente el desarrollo del determinante

$$\frac{1}{\left(\frac{\partial g}{\partial y}\right)^2}\det\begin{bmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} & -\frac{\partial g}{\partial z} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial x\partial z} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y\partial z} \\ -\frac{\partial g}{\partial z} & \frac{\partial^2 F}{\partial x\partial z} & \frac{\partial^2 F}{\partial y\partial z} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix}$$

Además, observamos que el signo de A queda determinado por el signo de

$$\left(\frac{\partial g}{\partial x}\right)^2\frac{\partial^2 F}{\partial y^2} - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 F}{\partial x\partial y} + \left(\frac{\partial g}{\partial y}\right)^2\frac{\partial^2 F}{\partial x^2} = -\det\begin{bmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x\partial y} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial y^2} \end{bmatrix}$$

que es justamente el negativo del determinante de la submatriz angular 3×3 que apareció en $B^2 - AC$.

La clave de este análisis está entonces en el determinante

$$H_F = \det\begin{bmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} & -\frac{\partial g}{\partial z} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial x\partial z} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 F}{\partial x\partial y} & \frac{\partial^2 F}{\partial y^2} & \frac{\partial^2 F}{\partial y\partial z} \\ -\frac{\partial g}{\partial z} & \frac{\partial^2 F}{\partial x\partial z} & \frac{\partial^2 F}{\partial y\partial z} & \frac{\partial^2 F}{\partial z^2} \end{bmatrix}$$

también llamado hessiano limitado de la matriz F, teniendo entonces que, si llamamos Δ_3 al determinante, la submatriz angular 3×3 del determinante H_F, entonces (ver Terorema 3.4.1)

(*) Si $H_F < 0$ y $\Delta_3 > 0$ (o sea, si $B^2 - AC < 0$ y $A < 0$) entonces la función f tendrá un máximo condicionado en el punto $\mathbf{p}$.

(*) Si $H_F < 0$ y $\Delta_3 < 0$ (o sea, si $B^2 - AC < 0$ y $A > 0$) entonces la función f tendrá un mínimo condicionado en el punto $\mathbf{p}$.

(*) Si $H_F > 0$ (o sea, si $B^2 - AC > 0$), entonces la función f no tiene extremo condicionado en el punto $\mathbf{p}$.

(*) Si $H_F = 0$ (o sea, si $B^2 - AC = 0$), no se puede concluir nada acerca de la naturaleza del punto $\mathbf{p}$.

Veamos un ejemplo.

Ejemplo 5. Estudiemos los extremos de la función $f(x, y, z) = x + y + z$ definida para $x > 0$, $y > 0$, $z > 0$ sujeta a la restricción $g(x, y, z) = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} - 1 = 0$. La función de Lagrange en este caso

$$F(x, y, z, \lambda) = x + y + z + \lambda\left(\frac{1}{x} + \frac{1}{y} + \frac{1}{z} - 1\right)$$

de donde

$$\frac{\partial F}{\partial x} = 1 - \frac{\lambda}{x^2} = 0, \qquad \frac{\partial F}{\partial y} = 1 - \frac{\lambda}{y^2} = 0,$$

$$\frac{\partial F}{\partial z} = 1 - \frac{\lambda}{z^2}, \qquad \frac{\partial F}{\partial \lambda} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} - 1 = 0$$

La única solución de este sistema es $x = y = z = 3$ (y $\lambda = 9$). Obtengamos los elementos del hessiano limitado de la función $F = f + \lambda g$.

$$\frac{\partial g}{\partial x} = -\frac{1}{x^2} \qquad \frac{\partial g}{\partial y} = -\frac{1}{y^2} \qquad \frac{\partial g}{\partial z} = -\frac{1}{z^2}$$

$$\frac{\partial^2 F}{\partial x^2} = \frac{2\lambda}{x^3} \qquad \frac{\partial^2 F}{\partial y^2} = \frac{2\lambda}{y^3} \qquad \frac{\partial^2 F}{\partial z^2} = \frac{2\lambda}{z^3}$$

$$\frac{\partial^2 F}{\partial x \partial y} = 0 \qquad \frac{\partial^2 F}{\partial x \partial z} = 0 \qquad \frac{\partial^2 F}{\partial y \partial z} = 0$$

de tal modo que en el punto (3, 3, 3) se tiene

$$H_F = \det \begin{bmatrix} 0 & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{2}{3} & 0 & 0 \\ \frac{1}{9} & 0 & \frac{2}{3} & 0 \\ \frac{1}{9} & 0 & 0 & \frac{2}{3} \end{bmatrix} = -\frac{4}{243} < 0 \qquad \Delta_3 = \det \begin{bmatrix} 0 & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{2}{3} & 0 \\ \frac{1}{9} & 0 & \frac{2}{3} \end{bmatrix} = -\frac{4}{243} < 0$$

Según el criterio establecido anteriormente, se tiene un mínimo condicionado de la función f en el punto (3, 3, 3). Obsérvese que lo que hemos establecido en este problema es que la suma de tres números positivos x, y, z, de modo que la suma de sus inversos sea la unidad, es un mínimo cuando los tres números son iguales a 3. ■

El análisis de las condiciones suficientes para extremos condicionados de funciones en los dos casos particulares anteriores nos presenta, de manera plausible, el siguiente criterio para el caso general de una función $y = f(x_1, x_2, \ldots, x_n)$ de n variables sujeta a la restricción $g(x_1, x_2, \ldots, x_n) = 0$, el cual no es más que la generalización natural de las condiciones establecidas en los casos previos.

Supongamos que f y g son funciones de clase $\mathscr{C}^2$ y que en $\mathbf{x}_0 \in U \subseteq \mathbb{R}^n$ (dominio de f), existe $\lambda \in \mathbb{R}$ tal que

$$\operatorname{grad} f(\mathbf{x}_0) + \lambda \operatorname{grad} g(\mathbf{x}_0) = 0$$

Consideramos entonces el hessiano limitado de la función de Lagrange $F = f + \lambda g$

$$H_F = \det \begin{bmatrix} 0 & -\frac{\partial g}{\partial x_1} & -\frac{\partial g}{\partial x_2} & \cdots & -\frac{\partial g}{\partial x_n} \\ -\frac{\partial g}{\partial x_1} & \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ -\frac{\partial g}{\partial x_2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \frac{\partial^2 F}{\partial x_2^2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n} \\ \vdots & & & & \\ -\frac{\partial g}{\partial x_n} & \frac{\partial^2 F}{\partial x_1 \partial x_n} & \frac{\partial^2 F}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{bmatrix}$$

donde todas las derivadas están evaluadas en $\mathbf{x}_0$. Sea

$$\Delta_{k+1} = \det \begin{bmatrix} 0 & -\frac{\partial g}{\partial x_1} & -\frac{\partial g}{\partial x_2} & \cdots & -\frac{\partial g}{\partial x_k} \\ -\frac{\partial g}{\partial x_1} & \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_k} \\ \vdots & & & & \\ -\frac{\partial g}{\partial x_k} & \frac{\partial^2 F}{\partial x_1 \partial x_k} & \frac{\partial^2 F}{\partial x_2 \partial x_k} & \cdots & \frac{\partial^2 F}{\partial x_k^2} \end{bmatrix}$$

$(2 \leq k \leq n)$. Entonces

a. Si $\Delta_k < 0$ para $k = 2, 3, \ldots, n$ la función f tiene un mínimo condicionado en $\mathbf{x}_0$.

b. Si $\Delta_3 > 0, \Delta_4 < 0, \Delta_5 > 0, \ldots$, la función f tiene un máximo condicionado en $\mathbf{x}_0$.

c. Si no todos los Δ_k son cero, pero no se tiene ninguna de las situaciones descritas en los dos incisos anteriores, entonces la función f no tiene extremo condicionado en $\mathbf{x}_0$.

Veamos un último ejemplo.

Ejemplo 6. Consideremos la función $f(x_1, x_2, \ldots, x_n) = x_1 x_2 \cdots x_n$ en donde $x_1, x_2, \ldots, x_n$ son positivos, sujeta a la restricción $g(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n - A = 0$ (A un número positivo dado). La función de Lagrange es

$$F(x_1, x_2, \ldots, x_n, \lambda) = x_1 x_2 \cdots x_n + \lambda(x_1 + x_2 + \cdots + x_n - A)$$

de donde

$$\frac{\partial F}{\partial x_1} = x_2 x_3 \cdots x_n + \lambda = 0 \qquad (1)$$

$$\frac{\partial F}{\partial x_2} = x_1 x_3 \cdots x_n + \lambda = 0 \qquad (2)$$

$$\vdots$$

$$\frac{\partial F}{\partial x_n} = x_1 x_2 \cdots x_{n-1} + \lambda = 0 \qquad (n)$$

$$\frac{\partial F}{\partial \lambda} = x_1 + x_2 + \cdots + x_n - A = 0 \qquad (n+1)$$

De las ecuaciones (1) a la (n), se obtiene $x_1 = x_2 = \ldots = x_n$, y de la ($n+1$) se obtiene $x_i = \frac{A}{n}$, $i = 1, 2, \ldots, n$. Por lo tanto sólo hay una solución al sistema, la cual corresponde al punto $\mathbf{p} = \left(\frac{A}{n}, \frac{A}{n}, \ldots, \frac{A}{n}\right) \in \mathbb{R}^n$. Los elementos del hessiano limitado de la función F son

$$\frac{\partial g}{\partial x_i} = 1, \quad i = 1, 2, \ldots, n$$

$$\frac{\partial^2 F}{\partial x_i^2} = 0, \quad \frac{\partial^2 F}{\partial x_i \partial x_j} = \underbrace{x_1 \cdots x_{i-1} x_{i+1} \cdots x_{j-1} x_{j+1} \cdots x_n}_{(n-2 \text{ factores})}$$

de modo que en el punto $\mathbf{p}$ se tiene

$$H_F = \det \begin{bmatrix} 0 & -1 & -1 & \ldots & -1 \\ -1 & 0 & C & \ldots & C \\ -1 & C & 0 & \ldots & C \\ \vdots & & & \ldots & \\ -1 & C & C & \ldots & 0 \end{bmatrix}$$

donde $C = \left(\frac{A}{n}\right)^{n-2}$. Se puede ver fácilmente que

$$\Delta_{k+1} = (-1)^k k C^{k-1} \qquad 2 \leq k \leq n$$

y por lo tanto los signos de estos determinates son

$$\Delta_3 > 0, \quad \Delta_4 < 0, \quad \Delta_5 > 0, \ldots$$

de modo que por el criterio establecido anteriormente, la función f tiene un máximo condicionado en el punto $\mathbf{p}$ (se puede llegar a esta conclusión razonando un poco en el contexto del problema; sin embargo nuestra intención aquí era obtenerla del criterio estudiado), que vale

$$f\left(\frac{A}{n}, \frac{A}{n}, \ldots, \frac{A}{n}\right) = \left(\frac{A}{n}\right)^n$$

Veamos la consecuencia de nuestro resultado. Hemos obtenido que $\left(\frac{A}{n}\right)^n$ es el valor más grande de la función $f(x_1, x_2, \ldots, x_n) = x_1 x_2 \cdots x_n$ sujeta a la restricción $x_1 + x_2 + \cdots + x_n = A$, es decir que

$$\left(\frac{A}{n}\right)^n \geq f(x_1, x_2, \ldots, x_n) = x_1 x_2 \cdots x_n$$

siempre que $x_1 + x_2 + \cdots + x_n = A$. Lo que podemos escribir con sólo la expresión

$$\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right)^n \geq x_1 x_2 \cdots x_n$$

o sea

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \ldots + x_n}{n}$$

donde $x_1, x_2, \ldots, x_n$ son cualesquiera números positivos. Así se obtiene la conclusión conocida de que la media geométrica de n números positivos nunca excede su media aritmética. ∎

Ejercicios (Capítulo 4, Sección 6)

En los ejercicios 1–6, compruebe, usando el Hessiano limitado, los resultados obtenidos sobre los extremos condicionados de las funciones $f(x, y)$ con las restricciones $g(x, y) = 0$ consideradas en los ejercicios indicados de la sección anterior.

1. Ejercicio 1: $f(x, y) = 2x - y$, $g(x, y) = 3x^2 + 2y^2 - 33/2 = 0$

2. Ejercicio 2: $f(x, y) = x + 3y$, $g(x, y) = 2x^2 + y^2 - 38 = 0$

3. Ejercicio 3: $f(x, y) = 3x - 7y$, $g(x, y) = 4x^2 + 3y^2 - 892/3 = 0$

4. Ejercicio 4: $f(x, y) = -2x - 5y$, $g(x, y) = 3x^2 + y^2 - 237 = 0$

5. Ejercicio 5: $f(x, y) = x + 2y$, $g(x, y) = x^2 + y^2 - 5 = 0$

6. Ejercicio 8: $f(x, y) = x^2y^2$, $g(x, y) = x^2 + y^2 - 1 = 0$

En los ejercicios 7–11, compruebe, usando el criterio establecido en esta sección (con los determinantes Δ_3 y H_F), los resultados obtenidos sobre los extremos condicionados de las funciones $f(x, y, z)$ con las restricciones $g(x, y, z) = 0$ consideradas en los ejercicios indicados de la sección anterior.

7. Ejercicio 16: $f(x, y, z) = 2x + 3y + z$, $g(x, y, z) = 4x^2 + 3y^2 + z^2 - 80 = 0$

8. Ejercicio 17: $f(x, y, z) = -x - 4y + 5z$, $g(x, y, z) = 5x^2 + y^2 + z^2 - 1030 = 0$

9. Ejercicio 18: $f(x, y, z) = x - y - z$, $g(x, y, z) = x^2 + 2y^2 + z^2 - 50 = 0$

10. Ejercicio 19: $f(x, y, z) = 2x - 3y + z$, $g(x, y, z) = 4x^2 + y^2 + z^2 - 704 = 0$

11. Ejercicio 20: $f(x, y, z) = x + z$, $g(x, y, z) = 2x^2 + 4y^2 + 5z^2 - 70 = 0$

12. En el ejercicio 23 de la sección anterior se obtuvieron los puntos $\mathbf{p}_1 = (5/9, 2/9, 2/9)$, $\mathbf{p}_2 = (-1, 1, 1)$ como candidatos a extremos condicionados para la función $f(x, y, z) = x^3 + 2y^2 + 2z^2 + x^2y + xy^2 + xz$, sujeta a la restricción $g(x, y, z) = x + y + z - 1 = 0$. Aplique los criterios estudiados en esta sección para concluír que la función f tiene en $\mathbf{p}_1$ un mínimo condicionado, y que en $\mathbf{p}_2$ no tiene extremo.

13. En el ejercicio 24 de la sección anterior se obtuvieron los puntos $\mathbf{p}_1 = (1/3, 1/3, 1/3)$, $\mathbf{p}_2 = (-1/3, 2/3, 2/3)$ como candidatos a extremos condicionados para la función $f(x, y, z) = x^3 + y^2 + z^2 + x^2y + xy^2 + xz$, sujeta a la restricción $g(x, y, z) = x + y + z - 1 = 0$. Aplique los criterios estudiados en esta sección para concluir que la función f tiene en $\mathbf{p}_1$ un mínimo condicionado, y que en $\mathbf{p}_2$ no tiene extremo.

14. En el ejercicio 31 de la sección anterior se consideró el problema de hallar la distancia mínima de los puntos de la hipérbola $x^2 - y^2 = 1$ al origen de coordenadas. Se obtuvieron dos puntos $\mathbf{p}_1 = (3/\sqrt{8}, 1/\sqrt{8})$, $\mathbf{p}_2 = (-3/\sqrt{8}, -1/\sqrt{8})$. Compruebe que estos dos puntos representan efectivamente mínimos para la función $f(x, y) = x^2 + y^2 =$ (cuadrado de la) distancia al origen del punto (x, y), sujeta a la restricción $g(x, y) = x^2 - y^2 - 1 = 0$.

(*) **15.** Considere la ecuación $\ln(x^2) + y^2 = 1$, la cual representa una curva en el plano. Use el método de los multiplicadores de Lagrange para determinar que los extremos condicionados de la función (cuadrado de la) distancia de (x, y) al origen, estando (x, y) en la curva dada, se pueden encontrar en los puntos $\mathbf{p}_1 = (1, 1)$, $\mathbf{p}_2 = (1, -1)$, $\mathbf{p}_3 = (-1, 1)$, $\mathbf{p}_4 = (-1, -1)$, $\mathbf{p}_5 = (\sqrt{e}, 0)$, $\mathbf{p}_6 = (-\sqrt{e}, 0)$. Compruebe que, de hecho, en los puntos $\mathbf{p}_{1,2,3,4}$ hay mínimos condicionados y que en $\mathbf{p}_{5,6}$ se dan máximos condicionados. Al ver la gráfica de la curva, es fácil convencerse de que los puntos $\mathbf{p}_{1,2,3,4}$ son los que en realidad se encuentran más cercanos al origen de coordenadas, pero que existen puntos de la curva que se encuentran más alejados del origen que los puntos $\mathbf{p}_{5,6}$. ¿Hay alguna contradicción en este hecho?

(*) **16.** En la discusión presentada en esta sección de la que se obtuvieron condiciones suficientes para que en el punto $\mathbf{p} = (x_0, y_0)$ exista un extremo condicionado de la función $z = f(x, y)$ sujeta a la restricción $g(x, y) = 0$, se arrancó del hecho de que $\frac{\partial g}{\partial y}(\mathbf{p}) \neq 0$. Suponga que esto no ocurre. Es decir, suponga que $\frac{\partial g}{\partial y}(\mathbf{p}) = 0$. Demuestre entonces que $\frac{\partial g}{\partial x}(\mathbf{p}) \neq 0$, y que en una vecindad de $\mathbf{p}$, la gráfica de $g(x, y) = 0$ puede verse como la gráfica de una función $x = \psi(y)$. Demuestre que, en tal caso, la función $z = f(x, y)$ sujeta a la restricción $g(x, y) = 0$ tiene un extremo condicionado en $\mathbf{p}$ si y solamente si la función $z = f(\psi(y), y)$ tiene un extremo local para $y = y_0$. Verifique que la segunda derivada de esta última función, en el punto $y = y_0$, se ve como

$$\frac{d^2z}{dy^2} = \frac{1}{\left(\frac{\partial g}{\partial x}\right)^2}\left[\frac{\partial^2 g}{\partial x\partial y}\frac{\partial g}{\partial x}\frac{\partial f}{\partial y} - \frac{\partial^2 g}{\partial x^2}\frac{\partial g}{\partial y}\frac{\partial f}{\partial y} + \left(\frac{\partial g}{\partial x}\right)^2\frac{\partial^2 f}{\partial y^2}\right.$$

$$\left. -2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 f}{\partial x\partial y} - \frac{\partial^2 g}{\partial y^2}\frac{\partial g}{\partial x}\frac{\partial f}{\partial x} + \frac{\partial^2 g}{\partial x\partial y}\frac{\partial g}{\partial y}\frac{\partial f}{\partial x} + \left(\frac{\partial g}{\partial y}\right)^2\frac{\partial^2 f}{\partial x^2}\right]$$

donde todas las derivadas parciales están evaluadas en $\mathbf{p}$ (se supone que f y g son funciones de clase $\mathscr{C}^2$, de modo que sus derivadas parciales mixtas son iguales). Demuestre finalmente que esta última expresión se escribe, en términos del Hessiano limitado definido en esta sección, como

$$\frac{d^2z}{dy^2}(y_0) = -\frac{H_F(x_0, y_0)}{\left(\frac{\partial g}{\partial x}(x_0, y_0)\right)^2}$$

Con ella, reobtenga *el mismo* criterio establecido en esta sección para concluír la existencia de un extremo condicionado de $z = f(x, y)$ con la restricción $g(x, y) = 0$, en el punto $\mathbf{p}$, en términos del Hessiano limitado $H_F(x_0, y_0)$.

Capítulo 5

Curvas en el espacio

En los tres capítulos anteriores hemos dedicado nuestro estudio a las funciones que, tomando valores reales, tienen su dominio en subconjuntos del espacio euclidiano n-dimensional $\mathbb{R}^n$. Estos tres capítulos cubren la parte "diferencial" del cálculo de estas funciones. La parte "integral" se abordará en el siguiente capítulo donde, por así decirlo, comienza la parte "integral" del presente libro. En alguna parte del estudio de funciones de varias variables tuvimos un acercamiento con las funciones más generales consideradas en este texto, del tipo $\mathbf{f}: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, cuyo dominio y codominio está en un espacio euclidiano de varias dimensiones. Esto fue simplemente para poner en un contexto global el estudio de la diferenciación de funciones compuestas. Sin embargo, podemos decir que lo que en realidad hemos estudiado hasta este momento es, como apuntábamos, el cálculo diferencial de las funciones $\mathbf{f}: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ en el caso particular $m = 1$. En este capítulo estudiaremos dichas funciones en el caso particular $n = 1$, dejando entonces que ahora el espacio euclidiano de varias dimensiones esté en el codominio de la función y que su dominio sea un subconjunto de $\mathbb{R}^1$ = la recta de los reales. Estas serán pues funciones del tipo $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, definidas en algún subconjunto I de $\mathbb{R}$ (digamos un intervalo) y tomando valores en el espacio $\mathbb{R}^n$. En realidad, concentraremos nuestro estudio en los casos $n = 2$ y $n = 3$ (el caso $n = 1$ fue estudiado en el primer curso de cálculo), ya que es en estos casos donde se cristaliza la riqueza geométrica del tema con visualizaciones concretas de los tópicos que se abordarán. Por supuesto que, entendiendo estos casos particulares, se llega a las generalizaciones correspondientes de un modo muy natural.
Comencemos.

5.1 Introducción. Límites y continuidad

Una función vectorial (o a valores vectoriales) de una variable real, es una función del tipo $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ la cual, a cada número real t de algún subconjunto I de $\mathbb{R}$ (el dominio de la función) le asocia un (y solamente un) valor $\mathbf{f}(t)$ en el espacio $\mathbb{R}^n$. (En todo este capítulo la variable independiente de f se denotará con la letra t o en algunas ocasiones, a partir de la sección 4, también con la letra s). Puesto que $\mathbf{f}(t)$ es un punto del espacio $\mathbb{R}^n$, éste tiene n coordenadas, las cuales son, en general, funciones de la variable t. Así, podemos escribir

$$\mathbf{f}(t) = (x_1(t), x_2(t), \ldots, x_n(t)) \in \mathbb{R}^n$$

donde $x_i: I \subseteq \mathbb{R} \to \mathbb{R}$, $i = 1, 2, \ldots, n$ son funciones reales de la variable real t, llamadas *funciones coordenadas* de la función f.

Como siempre, cuando solamente se proporciona la regla de correspondencia (como suele ocurrir) $\mathbf{f}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$, y no se hace explícito el dominio I de la función, se entiende que éste es el mayor subconjunto de la recta para el cual $x_i(t)$ hace sentido para toda $i = 1, 2, \ldots, n$ (el dominio natural de la función $\mathbf{f}$), es decir, el mayor subconjunto I de $\mathbb{R}$ para el cual $\mathbf{f}(t)$ hace sentido.

Esquemáticamente se tiene

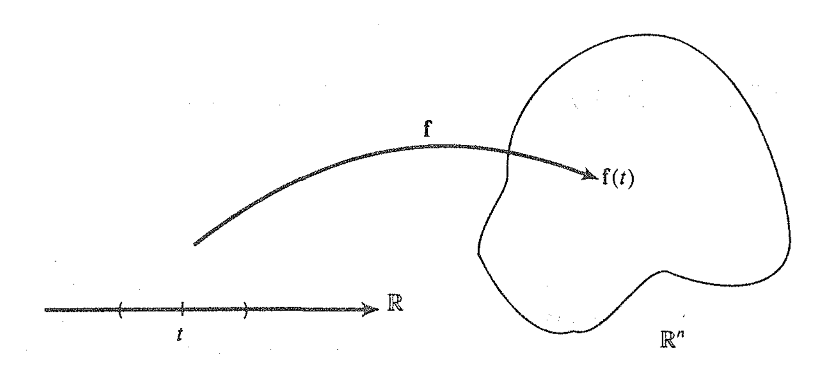


Figura 1. Una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$.

Las funciones $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ que nos interesa estudiar son las que son continuas en su dominio I. Para establecer el concepto de continuidad de estas funciones, antes tendremos que referirnos al concepto de límite.

La idea de límite para funciones $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es la misma que se estudió tanto en el primer curso de cálculo, como en el capítulo 2 para funciones de varias variables: si $t_0 \in I$, o t_0 es un punto frontera de I, se dirá que la función $\mathbf{f}$ tiende al límite $L \in \mathbb{R}^n$ cuando t tiende a t_0, si para valores de t cercanos a t_0 se tienen las correspondientes imágenes de $\mathbf{f}(t)$ cercanas a L. Para mayor precisión se tiene la definición siguiente

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ una función definida en un intervalo abierto I de $\mathbb{R}$ y sea t_0 un punto de I o un punto de frontera de I. Se dice que el límite de la función $\mathbf{f}$ cuando t tiende a t_0 es L ($\in \mathbb{R}^n$), lo cual se escribe como

$$\lim_{t \to t_0} \mathbf{f}(t) = \mathrm{L}$$

si dado cualquier $\epsilon > 0$ existe un $\delta > 0$ tal que

$$t \in I,\ 0 < |t - t_0| < \delta \Rightarrow \|\mathbf{f}(t) - \mathrm{L}\| < \epsilon$$

donde $\|\ \cdot\ \|$ es la norma euclideana de vectores de $\mathbb{R}^n$, es decir $\|\mathbf{x}\| = \left(\sum x_i^2\right)^{1/2}$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

Esquemáticamente (figura 2) ■

El siguiente teorema nos dice que el estudio de los límites de las funciones $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ está íntimamente relacionado con el estudio de los límites de las funciones coordenadas de $\mathbf{f}$.

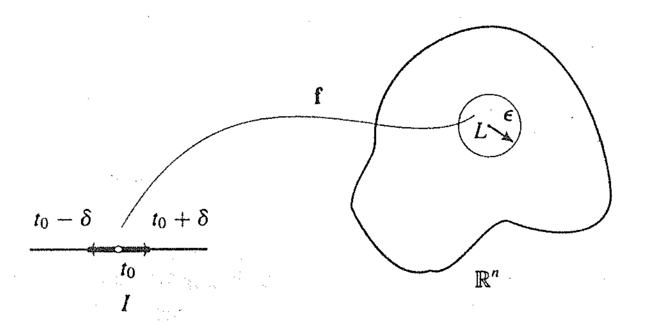


Figura 2. El límite de una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$.

Teorema 5.1.1 Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ una función como en la definición anterior. Entonces $\lim_{t \to t_0} \mathbf{f}(t) = \mathbf{L} = (\ell_1, \ell_2, \ldots, \ell_n) \in \mathbb{R}^n$ si y sólo si $\lim_{t \to t_0} x_i(t) = \ell_i$, donde $\mathbf{f}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. ■

Demostración. Supongamos primero que $\lim_{t \to t_0} \mathbf{f}(t) = \mathbf{L}$. Esto significa, según la definición dada anteriormente, que para cualquier $\epsilon > 0$ existe $\delta > 0$ de modo que si $t \in I$, $0 < |t - t_0| < \delta$ entonces $\|\mathbf{f}(t) - \mathbf{L}\| < \epsilon$. Pero como

$$\|\mathbf{f}(t) - \mathbf{L}\| = \|(x_1(t) - \ell_1, \ldots, x_n(t) - \ell_n)\| = \left(\sum_{i=1}^{n}(x_i(t) - \ell_i)^2\right)^{1/2} < \epsilon$$

se tiene que

$$|x_i(t) - \ell_i| \leq \left(\sum_{i=1}^{n}(x_i(t) - \ell_i)^2\right)^{1/2} < \epsilon$$

por lo tanto, dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$0 < |t - t_0| < \delta \Rightarrow |x_i(t) - \ell_i| < \epsilon$$

Lo que significa que $\lim_{t \to t_0} x_i(t) = \ell_i$, como se quería. Supongamos ahora que $\lim_{t \to t_0} x_i(t) = \ell_i$, $i = 1, 2, \ldots, n$. Esto quiere decir que para cualquier $\epsilon_i > 0$ hay un $\delta_i > 0$, tal que

$$t \in I,\ 0 < |t - t_0| < \delta_i \Rightarrow |x_i(t) - \ell_i| < \epsilon_i.$$

Sea $\epsilon > 0$ arbitrario y sea $\epsilon_i = \frac{\epsilon}{\sqrt{n}}$. Tome $\delta = \min(\delta_1, \delta_2, \ldots, \delta_n)$. Para tal δ se tiene

$$t \in I,\ 0 < |t - t_0| < \delta \Rightarrow |x_i(t) - \ell_i| < \frac{\epsilon}{\sqrt{n}} \quad \forall i = 1, 2, \ldots, n$$

Entonces

$$\|\mathbf{f}(t) - \mathbf{L}\| = \left(\sum_{i=1}^{n}(x_i(t) - \ell_i)^2\right)^{1/2} < \left(\sum_{i=1}^{n}\left(\frac{\epsilon}{\sqrt{n}}\right)^2\right)^{1/2} = \epsilon$$

Así se ha probado que para cualquier $\epsilon > 0$ existe $\delta > 0$ tal que

$$t \in I, 0 < |t - t_0| < \delta \Rightarrow \|\mathbf{f}(t) - \mathbf{L}\| < \epsilon$$

Esto significa que $\lim_{t \to t_0} \mathbf{f}(t) = \mathbf{L}$. Q.E.D.

Veamos un ejemplo.

Ejemplo 1. Sea $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ la función dada por $\mathbf{f}(t) = (t^2 + 1, 2t, t)$. Las funciones coordenadas de $\mathbf{f}$ son $x_1(t) = t^2 + 1$, $x_2(t) = 2t$, $x_3(t) = t$. Puesto que $\lim_{t \to t_0} x_1(t) = \lim_{t \to t_0}(t^2 + 1) = t_0^2 + 1$, $\lim_{t \to t_0} x_2(t) = \lim_{t \to t_0}(2t) = 2t_0$, $\lim_{t \to t_0} x_3(t) = \lim_{t \to t_0} t = t_0$, se tiene que, según el teorema anterior

$$\lim_{t \to t_0} \mathbf{f}(t) = \lim_{t \to t_0}(x_1(t), x_2(t), x_3(t)) = \left(\lim_{t \to t_0} x_1(t), \lim_{t \to t_0} x_2(t), \lim_{t \to t_0} x_3(t) \right) = (t_0^2 + 1, 2t_0, t_0)$$

Por ejemplo, si $t_0 = 1$, se tiene que $\lim_{t \to 1} \mathbf{f}(t) = (2, 2, 1)$. Esto significa que se puede tener a $\mathbf{f}(t)$ arbitrariamente cerca del punto $(2, 2, 1) \in \mathbb{R}^3$ con t suficientemente cerca de 1.

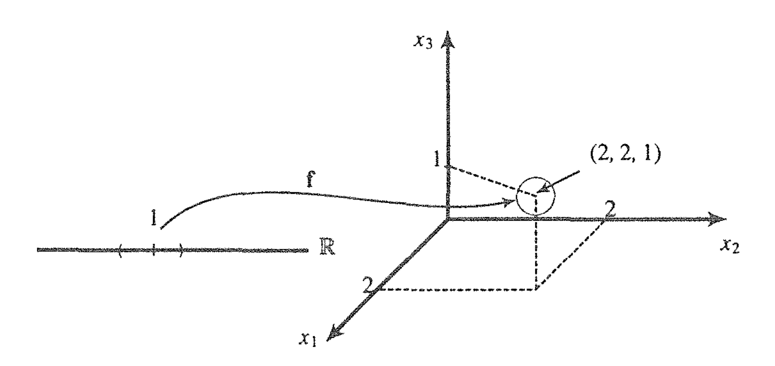


Figura 3. El límite de $\mathbf{f}(t) = (t^2 + 1, 2t, t)$ cuando $t \to 1$. ∎

En cierto sentido el comportamiento de la función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ lo determina completamente por el comportamiento de sus funciones coordenadas $x_i: I \subseteq \mathbb{R} \to \mathbb{R}$. Este hecho se empieza a sospecharse con el teorema anterior que nos dice que el límite de la función $\mathbf{f}$ está determinado por los límites de las funciones coordenadas x_i. Una situación análoga ocurrirá con la continuidad de la función $\mathbf{f}$; se verá que ésta es continua (con la definición que daremos a continuación) si y sólo si sus funciones coordenadas lo son.

Establezcamos la definición rigurosa de continuidad de este tipo de funciones.

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ una función definida en el subconjunto abierto I de $\mathbb{R}$ y sea $t_0 \in I$. Se dice que $\mathbf{f}$ es continua en t_0 si

$$\lim_{t \to t_0} \mathbf{f}(t) = \mathbf{f}(t_0)$$ ∎

Como siempre, lo que sugiere esta definición es que la propiedad de continuidad de una función $\mathbf{f}$ equivale a que movimientos pequeños de t en torno a t_0 producen movimientos pequeños en las imágenes $\mathbf{f}(t)$ alrededor de $\mathbf{f}(t_0)$; de otro modo, las imágenes cercanos a $\mathbf{f}(t_0)$ no tienen cambios bruscos para cambios pequeños del valor de t. Esto se puede ver mejor si escribimos la definición de continuidad dada anteriormente usando la definición de límite: la función $\mathbf{f}$ es continua en t_0 si dado $\epsilon > 0$ existe $\delta > 0$ de modo que

$$\underbrace{|t - t_0| < \delta}_{t \text{ está cerca de } t_0 \text{ en menos que } \delta} \quad \underbrace{\Rightarrow}_{\text{entonces}} \quad \underbrace{\|\mathbf{f}(t) - \mathbf{f}(t_0)\| < \epsilon}_{\mathbf{f}(t) \text{ está cerca de } \mathbf{f}(t_0) \text{ en menos que } \epsilon}$$

(la condición $t \neq t_0$ que en la definición de límite estaba escrita como $|t - t_0| > 0$, ahora queda eliminada pues la definición de continuidad requiere que $\mathbf{f}(t_0)$ exista).

La cadena de igualdades

$$\begin{aligned} \lim_{t \to t_0} \mathbf{f}(t) &= \lim_{t \to t_0} (x_1(t), x_2(t), \ldots, x_n(t)) \\ &= \left(\lim_{t \to t_0} x_1(t), \lim_{t \to t_0} x_2(t), \ldots, \lim_{t \to t_0} x_n(t)\right) \\ &= \left(x_1(t_0), x_2(t_0), \ldots, x_n(t_0)\right) \\ &= \mathbf{f}(t_0) \end{aligned}$$

junto con algunos comentarios pertinentes por parte del lector, (aplicando desde luego el teorema demostrado previamente y la definición de continuidad de $\mathbf{f}$ y de sus funciones coordenadas), deberá sin duda proporcionar el argumento contundente que pruebe la validez del teorema siguiente.

Teorema 5.1.2 Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ una función definida en el intervalo abierto I de $\mathbb{R}$, digamos que $\mathbf{f}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. Sea $t_0 \in I$. La función $\mathbf{f}$ es continua en t_0 si y sólo si sus funciones coordenadas $x_i: I \subseteq \mathbb{R} \to \mathbb{R}$ lo son. ■

Una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ definida en el intervalo abierto I de $\mathbb{R}$ se dice ser continua en I (o simplemente continua) si lo es en todo $t \in I$.

Ejemplo 2. La función $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$ dada por

$$\mathbf{f}(t) = (t^3 + 1, t^2 - 2t + 1)$$

es continua, pues sus funciones coordenadas son polinomiales y, por tanto, continuas. ■

Ejemplo 3. La función $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dada por

$$\mathbf{f}(t) = \begin{cases} \left(t, t^2, \dfrac{\operatorname{sen} t}{t}\right) & \text{si } t \neq 0 \\ (0, 0, 0) & \text{si } t = 0 \end{cases}$$

es discontinua en $t = 0$ pues

$$\lim_{t \to t_0} \mathbf{f}(t) = \lim_{t \to 0} \left(t, t^2, \frac{\operatorname{sen} t}{t}\right) = (0, 0, 1) \neq (0, 0, 0) = \mathbf{f}(0)$$ ■

La continuidad de una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ definida en el intervalo I de $\mathbb{R}$ es el ingrediente esencial que nos permite definir el tipo de objetos de nuestro estudio en este capítulo.

Ejercicios (Capítulo 5, Sección 1)

1. Sea $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ la función $\mathbf{f}(t) = (t^2, 3t - 1, t^3 - 2)$. Determine:

a. $\mathbf{f}(1)$;

b. $\mathbf{f}(3) - \mathbf{f}(2)$;

c.. $0.5\mathbf{f}(-1)$;

d.. $\mathbf{f}(1) \cdot \mathbf{f}(3)$;

e.. $\|\mathbf{f}(2)\|\mathbf{f}(1)$;

f.. $\mathbf{f}(1) \times \mathbf{f}(-1)$;

g.. $\mathbf{f}(0) \cdot \mathbf{f}(1) \times \mathbf{f}(2)$

2. Sean $f_1: I \subseteq \mathbb{R} \to \mathbb{R}$ y $f_2: J \subseteq \mathbb{R} \to \mathbb{R}$ dos funciones reales definidas en los subconjuntos I y J de $\mathbb{R}$. ¿Cuál es el dominio (natural) U de la función $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (f_1(t), f_2(t))$? En general, si $f_i: I_i \to \mathbb{R}$ es una función definida en el subconjunto I_i de $\mathbb{R}$, $i = 1, 2, \ldots, k$, ¿cuál es el dominio U de la función $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^k$, $\mathbf{f}(t) = (f_1(t), f_2(t), \ldots, f_k(t))$?

En los ejercicios 3–10 determine el dominio U de las funciones siguientes

3. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, 1)$

4. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (\sqrt{t}, t^2)$

5. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (\sqrt{1+t}, \sqrt{1-t})$

6. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (1/t, \ln t)$

7. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t, t^2, t^3)$

8. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (\ln(t^2 + t + 1), \ln(t^2 - t + 1), \ln(t^2 + 1))$

9. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t, \operatorname{arcsen} t, \ln t)$

10. $\mathbf{f}: U \subseteq \mathbb{R} \to \mathbb{R}^4$, $\mathbf{f}(t) = (\operatorname{arcsen} t, \operatorname{arcsen} 2t, \operatorname{arcsen} 3t, \operatorname{arcsen} 4t)$

11. Se sabe que $\lim_{t\to 2}(t, t) = (2, 2)$. Puesto que $\epsilon > 0$, determine el $\delta > 0$ que verifique la validez de este límite.

12. Repita el ejercicio anterior con el límite $\lim_{t\to 2}(3t, -t) = (6, -2)$.

13. Se dice que la función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es una *función constante*, si $\mathbf{f}(t_1) = \mathbf{f}(t_2)\ \forall t_1, t_2 \in I$. En tal caso se escribe $\mathbf{f}(t) = \mathbf{c}$ (donde $\mathbf{c} \in \mathbb{R}^n$ es igual $\mathbf{f}(t)$ para cualquier $t \in I$). Por ejemplo, $\mathbf{f}(t) = (1, 3)$ y $\mathbf{f}(t) = (2, 0, -4)$ son funciones constantes. Enuncie con precisión y demuestre el resultado siguiente: el límite de una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ constante es igual al valor de la constante.

En los ejercicios 14–20, calcule los límites indicados

14. $\lim_{t\to 1}(3t - 1, t)$

15. $\lim_{t\to 1}\left(\frac{t^2-1}{t-1}, \frac{t^2-1}{t+1}\right)$

16. $\lim_{t\to 0}\left(\frac{\operatorname{sen} t}{t}, \frac{t}{\cos t}\right)$

17. $\lim_{t\to 0}\left(\frac{1-\cos t}{t^2}, \frac{1-\cos 2t}{t^2}\right)$

18. $\lim_{t\to 1}\left(\frac{t^2-1}{t-1}, \frac{t^3-1}{t^2-1}, \frac{t^4-1}{t^3-1}\right)$

19. $\lim_{t\to 0}\left(\frac{(x+t)^2-x^2}{t}, \frac{(x+t)^3-x^3}{t}, \frac{(x+t)^4-x^4}{t}\right)$

20. $\lim_{t\to 0}\left(\frac{\operatorname{sen} 2t}{\operatorname{sen} 3t}, \frac{\cos 2t}{\cos 3t}, \frac{\operatorname{sen} 4t}{\cos 5t}\right)$

En los ejercicios 21–25, exponga la continuidad de las funciones

21. $\mathbf{f}(t) = \left(\frac{t^2-1}{t-1}, \frac{t^2-1}{t+1}\right)$

22. $\mathbf{f}(t) = (t, \operatorname{sgn} t)$

23. $\mathbf{f}(t) = (\operatorname{sgn}(t-1), \operatorname{sgn}(t+1))$

24. $\mathbf{f}(t) = \begin{cases} \left(t, \frac{\operatorname{sen} t}{t}\right) & \text{si } t \neq 0 \\ (0, 1) & \text{si } t = 0 \end{cases}$

25. $\mathbf{f}(t) = \begin{cases} \left(\frac{\operatorname{sen} 2t}{\operatorname{sen} 3t}, \frac{\cos 2t}{\cos 3t}, \frac{\operatorname{sen} 4t}{\cos 5t}\right) & \text{si } t \neq 0 \\ (2/3, 1, 0) & \text{si } t = 0 \end{cases}$

26. Suponga que la función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es continua en I. Demuestre que la función $\varphi: I \to \mathbb{R}$, $\varphi(t) = \|\mathbf{f}(t)\|$ es continua en I.

27. Si $\mathbf{f}, \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ son dos funciones continuas, demuestre que las siguientes funciones son continuas;

a. $\mathbf{f} + \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, $(\mathbf{f} + \mathbf{g})(t) = \mathbf{f}(t) + \mathbf{g}(t)$

b. $\mathbf{f} \cdot \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, $(\mathbf{f} \cdot \mathbf{g})(t) = \mathbf{f}(t) \cdot \mathbf{g}(t)$

c. $\mathbf{f} \times \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, $(\mathbf{f} \times \mathbf{g})(t) = \mathbf{f}(t) \times \mathbf{g}(t)$

28. Por medio de ejemplos concretos, muestre que la afirmación recíproca del ejercicio anterior es, en general, falsa. Es decir, dé ejemplos de funciones $\mathbf{f}, \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ por los que $\mathbf{f} + \mathbf{g}$, $\mathbf{f} \cdot \mathbf{g}$, $\mathbf{f} \times \mathbf{g}$ son funciones continuas, sin que $\mathbf{f}$ y/o $\mathbf{g}$ sean continuas.

5.2 Caminos en $\mathbb{R}^n$. Consideraciones y ejemplos preliminares

La definición más importante del capítulo, punto que nos presenta los objetos a estudiar en él, es la siguiente

Definición. Una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ continua, definida en el intervalo I de $\mathbb{R}$, se llama *camino* o *trayectoria* en el espacio $\mathbb{R}^n$ ...
[Antes de completar la definición haremos un par de comentarios. En ella nada se dice de la naturaleza del intervalo $I \subseteq \mathbb{R}$ donde está definida $\mathbf{f}$. Si éste es abierto, ya sabemos qué significa que la función $\mathbf{f}$ sea continua en I. Si no lo es, digamos por ejemplo que sea cerrado $I = [a, b]$, entonces decimos que $\mathbf{f}$ es continua en I si lo es en (a, b) y si

$$\lim_{t \to a^-} \mathbf{f}(t) = \mathbf{f}(a) \quad y \quad \lim_{t \to b^+} \mathbf{f}(t) = \mathbf{f}(b)$$

Los caminos de $\mathbb{R}^n$ son los objetos matemáticos de nuestro estudio en este capítulo. Ellos son pues, según la definición anterior, *funciones continuas* $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ definidas en el intervalo I de $\mathbb{R}$ y tomando valores en el espacio $\mathbb{R}^n$.]
... Si la función continua $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ está definida en el intervalo cerrado $I = [a, b]$, diremos que el punto $\mathbf{f}(a) \in \mathbb{R}^n$ es el *punto inicial* del camino o trayectoria $\mathbf{f}$, en tanto que $\mathbf{f}(b) \in \mathbb{R}^n$ es el *punto final* de él. Si acontece que $\mathbf{f}(a) = \mathbf{f}(b)$, diremos que el camino $\mathbf{f}$ es *cerrado*. Si la función $\mathbf{f}$ es inyectiva en I (es decir, $\forall t_1, t_2 \in I, t_1 \neq t_2 \Rightarrow \mathbf{f}(t_1) \neq \mathbf{f}(t_2)$), diremos que $\mathbf{f}$ es un *camino simple*. Si se tiene que $\mathbf{f}(a) = \mathbf{f}(b)$ y la función $\mathbf{f}$ restringida al intervalo $[a, b)$ es inyectiva, diremos que el camino es *cerrado simple*. ■

Nos inclinaríamos ahora a tener una visualización geométrica de este tipo de objetos recién definidos, lo cual nos lleva de manera natural a pensar en "la gráfica" de un camino. Ciertamente no hemos definido lo que se debe entender por gráfica de una función como las que estamos estudiando. Sin embargo, no resultará difícil establecer tal definición si vemos de nuevo cómo es este concepto en el caso de las funciones reales de una variable real y el caso de funciones de varias variables (ver sección 2, capítulo 2). Haciendo una generalización (o adecuación) natural para el caso de funciones $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, diremos que la gráfica de una función como ésta es el conjunto

$$\{(t, \mathbf{f}(t)) \in \mathbb{R}^{n+1} | t \in I\} \subset \mathbb{R}^{n+1}$$

(abusando un poco de la notación, hemos puesto en la segunda coordenada de $(t, \mathbf{f}(t))$, el punto $\mathbf{f}(t) \in \mathbb{R}^n$). Sugerimos al lector que intente visualizar este concepto en el caso $n = 2$ (único caso en que se puede hacer, para tener así los puntos de la gráfica de $\mathbf{f}$ en el espacio $\mathbb{R}^3$), para que vea y sienta las dificultades que se presentan. Este es, pues, un concepto que en nada nos ayuda a tener representaciones geométricas "decentes" de los nuevos objetos estudiados.

Sin embargo, sí hay una manera de "ver" nuestros caminos $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, que nos ayudará mucho a entender su comportamiento. La manera correcta de abordar la geometría de estos objetos matemáticos es por medio del siguiente concepto.

Definición. Se llama *traza* del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, al conjunto (subconjunto de $\mathbb{R}^n$) de las imágenes de $\mathbf{f}$; es decir

$$\text{traza de } \mathbf{f} = \{\mathbf{f}(t) \in \mathbb{R}^n | t \in I\} \subset \mathbb{R}^n$$ ■

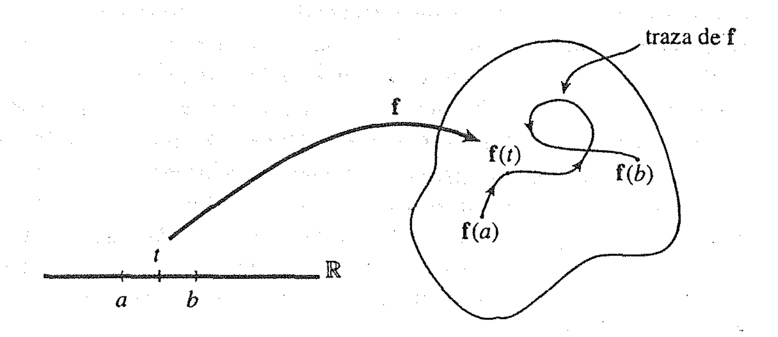


Figura 1. Traza de un camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$.

Así, la traza de un camino **f** es algo así como la "huella" de las imágenes de **f**. (Esta huella se deja en el espacio donde viven las imágenes de la función).

Conviene siempre tener en mente que la variable independiente t recorre de izquierda a derecha el intervalo I e imaginar a las imágenes correspondientes moviéndose en el espacio $\mathbb{R}^n$, dejando así marcada *de manera continua* su traza. En realidad, esta visión "mecánica" de los caminos es muy importante para entender adecuadamente algunos puntos en el estudio que emprenderemos en este capítulo. Aún más, se usará la letra t para designar a la variable independiente de nuestros caminos, y precisamente para que ésta sea pensada como *tiempo*. Ante esta perspectiva, podemos pensar que cada camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$ es como "la fórmula matemática de una carretera en el espacio", la cual, como veremos, tiene dos ingredientes importantes que la caracterizan: la forma de la carretera en sí y la manera como ésta es recorrida.

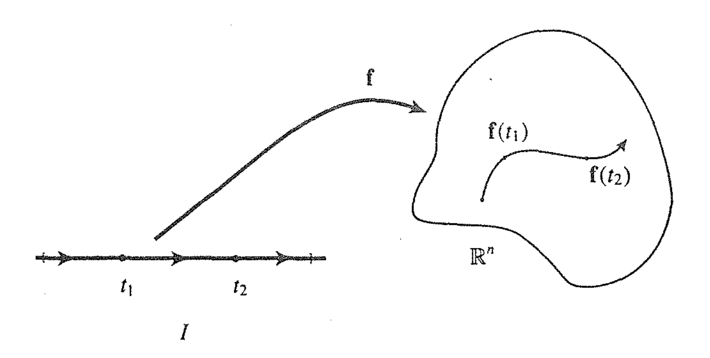


Figura 2. Diagrama de la traza de una función.

Designaremos con la palabra *curva* (en $\mathbb{R}^n$) la traza de un camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$. Esta palabra tiene un contenido geométrico más fuerte que "camino" o "trayectoria", lo cual resulta muy conveniente para nuestros propósitos, pues en ocasiones debemos distinguir al camino en sí, el cual es *una función* continua $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$, del aspecto visual que presentan sus imágenes en conjunto, es decir, de

la curva asociada a $\mathbf{f}$. Sin embargo, cuando no haya peligro de confusión, nos podremos referir a propiedades del camino $\mathbf{f}$ como propiedades de la curva asociada a $\mathbf{f}$.

Pensemos por ejemplo en un camino en $\mathbb{R}^2$, digamos $\mathbf{f}: [a, b] \to \mathbb{R}^2$, $\mathbf{f}(t) = (x(t), y(t))$. El punto inicial de él es $\mathbf{f}(a) = (x(a), y(a))$ y el punto final es $\mathbf{f}(b) = (x(b), y(b))$. Supongamos que $\mathbf{f}$ es una función inyectiva. En tal caso, siendo el camino $\mathbf{f}$ simple, podremos decir que *la curva* $\{\mathbf{f}(t) \in \mathbb{R}^2 | t \in [a, b]\}$ *es* (una curva) *simple*. Veamos de cerca cuál es la implicación geométrica de este concepto. Supongamos que la curva pasa por el punto (x_0, y_0) (es decir, para algún $t_0 \in I$ se tiene $\mathbf{f}(t_0) = (x_0, y_0)$). El hecho de ser la función $\mathbf{f}$ inyectiva significa que la curva *no vuelve a pasar por* (x_0, y_0) (si así lo hiciera, existiría $t_1 \in I$ tal que $\mathbf{f}(t_1) = (x_0, y_0) = \mathbf{f}(t_0)$, lo que viola la propiedad de inyectividad de $\mathbf{f}$). Por lo tanto, una curva simple es una curva que no tiene autointersecciones. Por ejemplo, los siguientes caminos (las siguientes curvas) en $\mathbb{R}^2$ no son simples

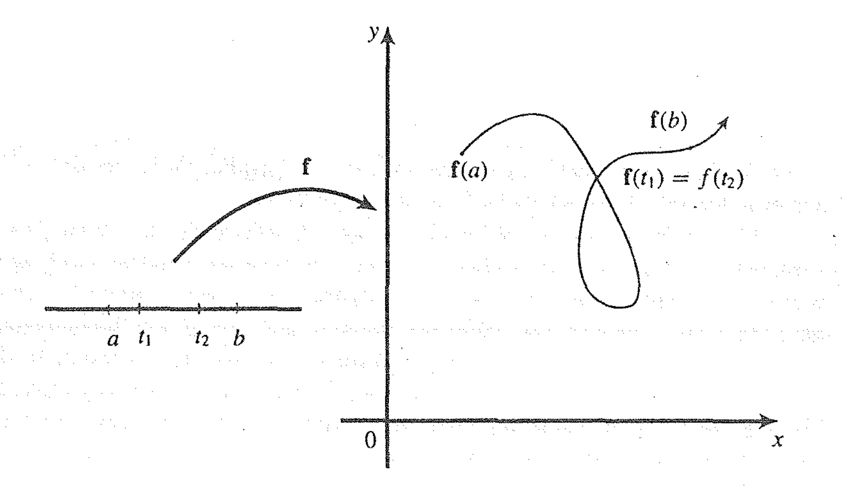


Figura 3. Curvas que no son simples.

En el caso de caminos cerrados simples $\mathbf{f}: [a, b] \to \mathbb{R}^2$ (que no equivale a que sean cerrados y simples, —pues de hecho tales caminos... ¡no existen!— ¿por qué?) el aspecto geométrico es el siguiente

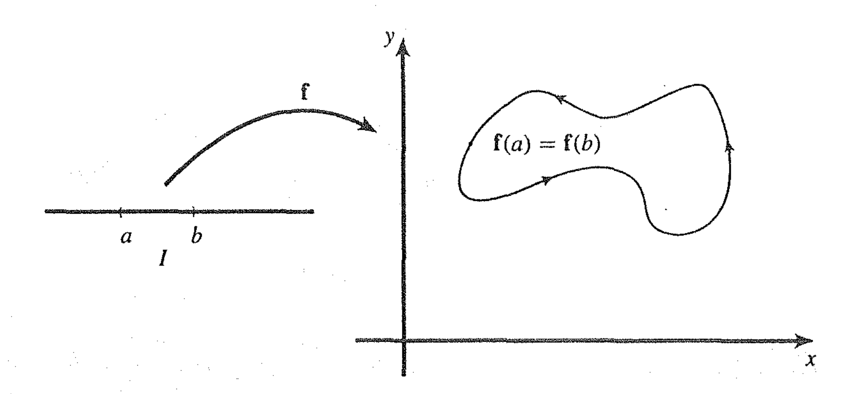


Figura 4. Curvas cerradas simples.

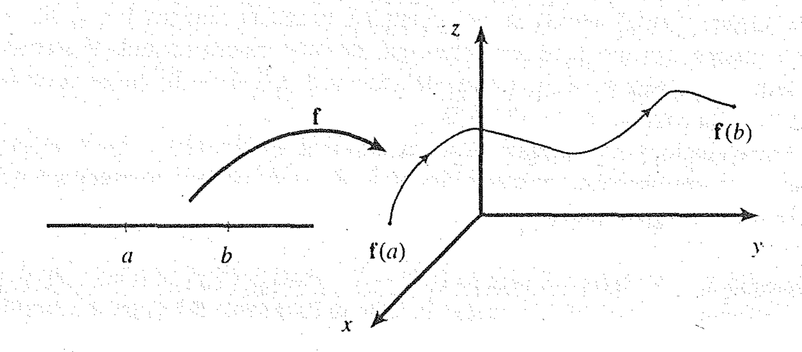


Figura 5. Curvas simples.

Del mismo modo, curvas simples en $\mathbb{R}^3$ tienen aspectos como los que se muestran en la figura 5. Veamos algunos ejemplos concretos.

Ejemplo 1. Consideremos la función continua $\varphi: \mathbb{R} \to \mathbb{R}$. Esta es una función real de una variable real $x \in I$. Como sabemos, la gráfica de esta función es (por definición) el conjunto

$$\{(x, y)|x \in I, y = \varphi(x)\} \subset \mathbb{R}^2$$

al cual siempre nos hemos referido como "una curva" en el plano.¿ Es ésta una "curva" en el sentido que estamos manejando en el presente capítulo?, es decir, ¿es la gráfica de la función $y = \varphi(x)$ igual a la traza de una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$? Nótese que si definimos la función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ (I es el mismo intervalo de la recta donde está definida φ) como

$$\mathbf{f}(t) = (t, \varphi(t))$$

tenemos que $\mathbf{f}$ es continua (pues sus funciones coordenadas $x(t) = t$, $y(t) = \varphi(t)$ lo son) y además

$$\text{traza de } \mathbf{f} = \{(t, \varphi(t))|t \in I\} = \text{gráfica de } \varphi$$

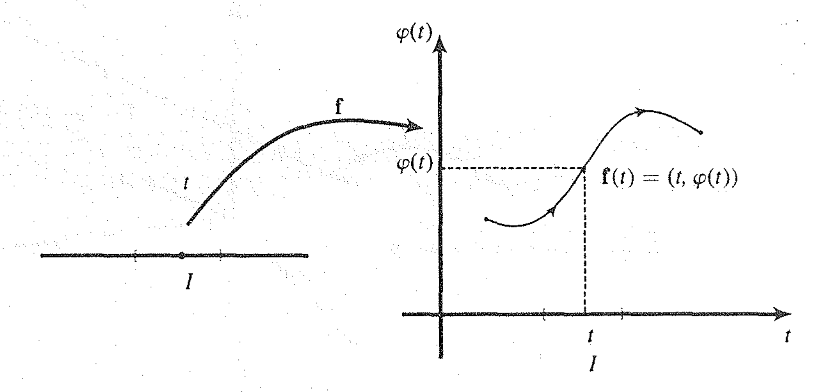


Figura 6. Gráfica de la función φ. ■

De este modo tenemos pues que las curvas que son gráficas de funciones continuas $\varphi: I \subseteq \mathbb{R} \to \mathbb{R}$, son "curvas" (en el sentido de este capítulo), trazas de caminos $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, \varphi(t))$. Sin embargo, es claro que nuestro concepto de curva en este capítulo abarca situaciones mucho más generales que ésta; es decir, la traza del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ no es necesariamente la gráfica de una función continua $\varphi: I \subseteq \mathbb{R} \to \mathbb{R}$.

Por ejemplo, la curva asociada al camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^2)$ es la gráfica de la función $\varphi(x) = x^2$, sin embargo, cualquier curva $\mathbf{f}: [a, b] \to \mathbb{R}^2$ cerrada no puede ser la gráfica de una función $\varphi: [a, b] \to \mathbb{R}$ (¿por qué?).

Ejemplo 2. La traza del camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, dada por $\mathbf{f}(t) = (t, mt + b)$ es una recta en el plano, de pendiente m y ordenada al origen b. Esto se desprende del ejemplo anterior

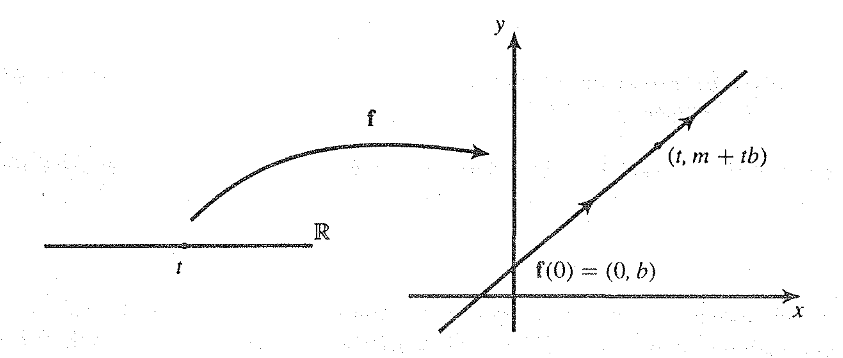


Figura 7. Traza del camino $\mathbf{f}(t) = (t, mt + b)$.

Incluso, la traza del camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, dada por $\mathbf{f}(t) = (at + x_0, bt + y_0, ct + z_0)$ es una recta en el espacio que pasa por el punto (x_0, y_0, z_0) (correspondiente a $t = 0$), y que es paralela al vector $\mathbf{v} = (a, b, c)$. Esto es un hecho que se desprende de las ecuaciones $x = at + x_0$, $y = bt + y_0$, $z = ct + z_0$, que conocemos como ecuaciones paramétricas de una recta en el espacio

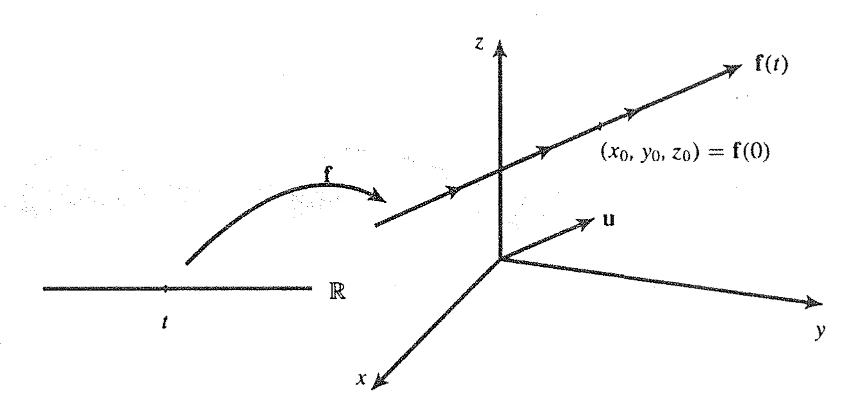


Figura 8. Una recta en $\mathbb{R}^3$, traza del camino $\mathbf{f}(t) = (at + x_0, bt + y_0, ct + z_0)$. ■

Ejemplo 3. Consideremos el camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (r\cos t, r\,\text{sen}\, t)$. Obsérvese que las funciones coordenadas $x(t) = r\cos t$, $y(t) = r\,\text{sen}\, t$ cumplen con

$$(x(t))^2 + (y(t))^2 = (r\cos t)^2 + (r\,\text{sen}\, t)^2 = r^2 \qquad \forall t \in [0, 2\pi]$$

de modo que en cualquier instante $t \in [0, 2\pi]$, el punto $\mathbf{f}(t)$ se encuentra en el círculo de centro en el origen y radio r, $x^2 + y^2 = r^2$. Así, la traza del camino $\mathbf{f}$ es un círculo de centro en el origen y radio r, cuyo punto inicial es $\mathbf{f}(0) = (r\cos 0, r\,\text{sen}\, 0) = (r, 0)$, y cuyo punto final es $\mathbf{f}(2\pi) = (r\cos 2\pi, r\,\text{sen}\, 2\pi) = (r, 0) = \mathbf{f}(0)$; es decir, se trata de una curva cerrada simple (¿por qué?) que se recorre partiendo del punto $(r, 0)$, en sentido contrario a las manecillas del reloj (antihorario).

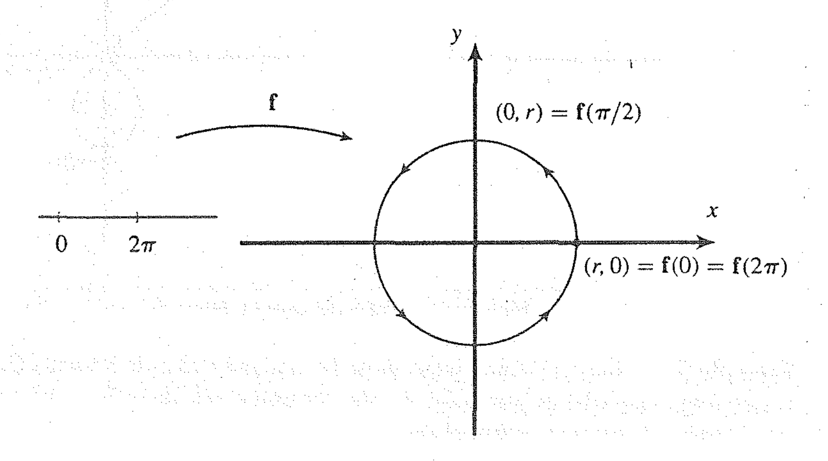


Figura 9. La traza del camino cerrado simple $\mathbf{f}(t) = (r\cos t, r\,\text{sen}\, t)$, $t \in [0, 2\pi]$.

De hecho, en este caso podemos ver la variable independiente t como el valor del ángulo (medido en radianes) que forma el vector $\mathbf{f}(t)$ con la parte positiva del eje x.

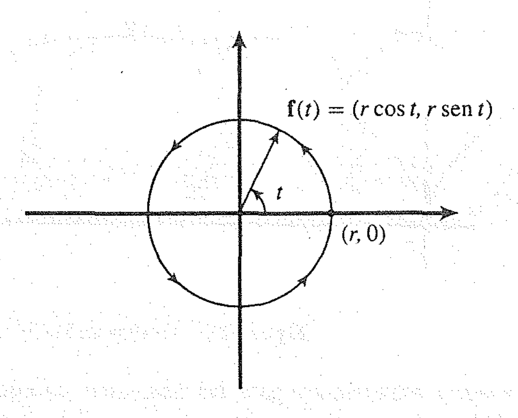


Figura 10. Interpretación geométrica de la variable t en el camino $\mathbf{f}(t) = (r\cos t, r\,\text{sen}\, t)$.

■

Ejemplo 4. Considere el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, dado por $\mathbf{f}(t) = (t^3 - t, t^2 - 1)$. Obsérvese que $\mathbf{f}(-1) = (0, 0) = \mathbf{f}(1)$, de modo que este es un camino que no es simple. Podríamos "rastrear" su traza con algunos valores de t, por ejemplo, $\mathbf{f}(-2) = (-6, 3)$, $\mathbf{f}(-1) = (0, 0)$, $\mathbf{f}(0) = (0, -1)$, $\mathbf{f}(1) = (0, 0)$, $\mathbf{f}(2) = (6, 3), \ldots$ Estos valores, acompañados de alguna reflexión adicional sobre los posibles valores de $x(t)$, $y(t)$ para diferentes valores de t, nos pueden dar una idea de la forma geométrica de la curva.

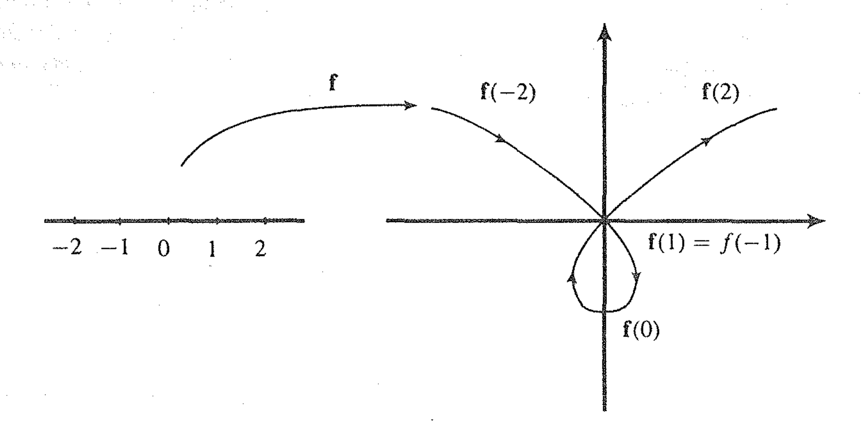


Figura 11. Traza del camino $\mathbf{f}(t) = (t^3 - t, t^2 - 1)$. ■

Ejemplo 5. Una curva muy importante en matemáticas es la llamada CICLOIDE, definida como la curva que describe un punto fijo de una circunferencia de radio a, la cual rueda, sin deslizarse, a una velocidad constante sobre el eje x.

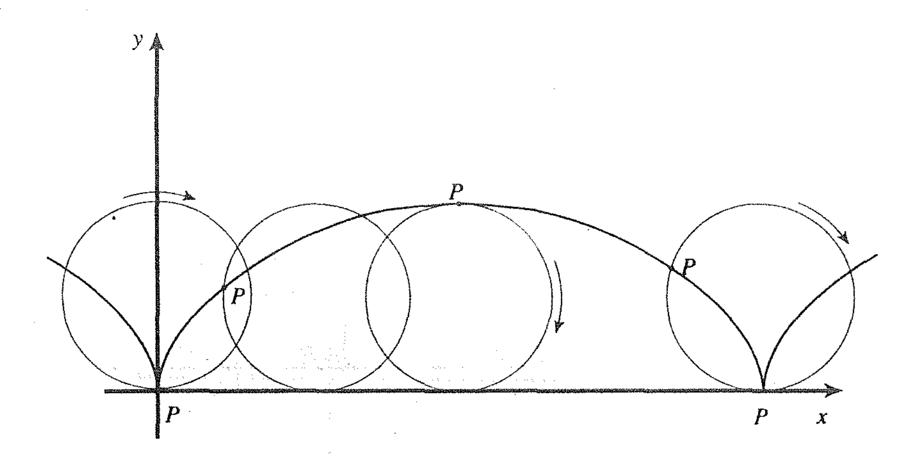


Figura 12. Gráfica de la CICLOIDE.

Vamos a conseguir expresiones para las funciones coordenadas x, $y: \mathbb{R} \to \mathbb{R}$, de un camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (x(t), y(t))$ que tenga por traza una cicloide. En un instante dado, el círculo que está rodando se ve como

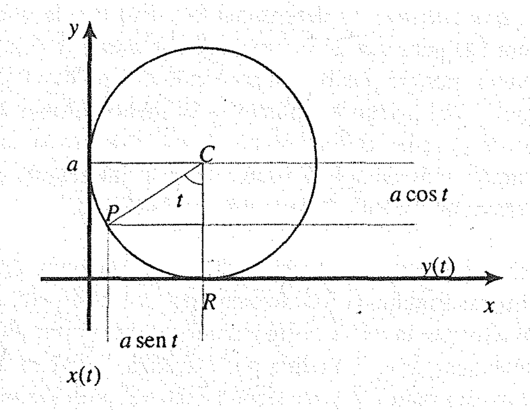


Figura 13. El círculo que rueda para describir la cicloide.

Sea t el ángulo, en radianes, que forma (la línea que contiene) el radio $\overline{\mathrm{CP}}$ con la parte negativa del eje y, como se muestra en la figura 13.

Nótese que la distancia $|\overline{\mathrm{OR}}|$ es justamente la longitud del arco de circunferencia RP, que es igual a at, de modo que el punto C tiene por coordenadas (at, a). Si $(x(t), y(t))$ denotan las coordenadas del punto $\mathbf{p}$, se tienen, atendiendo a la figura 13, las siguientes relaciones

$$y(t) = a - a\cos t = a(1 - \cos t)$$
$$x(t) = at - a\,\mathrm{sen}\,t = a(t - \mathrm{sen}\,t)$$

Así, el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$ dado por

$$\mathbf{f}(t) = (a(t - \mathrm{sen}\,t), a(1 - \cos t))$$

tiene por traza una cicloide.

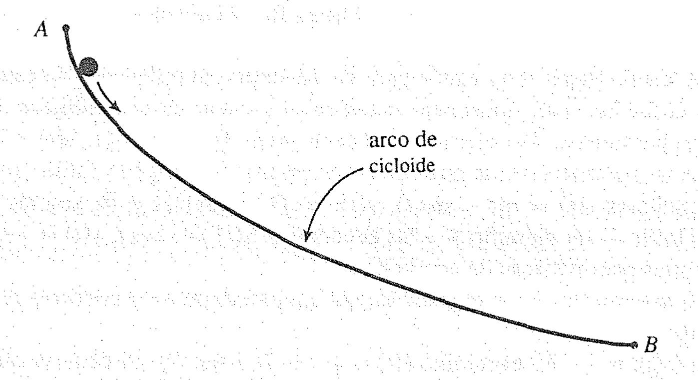


Figura 14. Un cuerpo cayendo de A a B.

La Cicloide es una curva conocida por los matemáticos desde el siglo XVIII. Surgió como respuesta a un famoso problema planteado por Johann Bernoulli en 1691: el "problema de la

braquistócrona", que consiste en determinar la curva con la que se deben unir los puntos A y B (como en la figura 14) para que un cuerpo que cae desde A llegue a B, siguiendo la trayectoria de la curva, en el menor tiempo posible (suponiendo que no hay fricción). Esta curva lleva el nombre de "braquistócrona" (del griego brachistos = el menor, cronos = tiempo) y fue el mismo Johann Bernoulli (entre otros) quien descubrió que la cicloide era la curva que solucionaba al problema. (Para una exposición más completa acerca de este interesante problema, se puede consultar por ejemplo, el apéndice del capítulo 3 de la referencia [PiII]).

Ejemplo 6. Consideremos el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por $\mathbf{f}(t) = (r\cos t, r\,\mathrm{sen}\,t, t)$. Obsérvese cómo las dos primeras funciones coordenadas $x(t) = r\cos t$, $y(t) = r\,\mathrm{sen}\,t$ satisfacen $x(t)^2 + y(t)^2 = r^2, \forall t \in \mathbb{R}$, de modo que la curva correspondiente debe estar dibujada en el cilindro circular recto $x^2 + y^2 = r^2$. Por otra parte, a medida que t avanza, la tercera función coordenada de $\mathbf{f}$, $z(t) = t$, que marca la altura del punto $\mathbf{f}(t)$, va siendo cada vez más grande. En base a estas condiciones, se puede ver que la curva que describe $\mathbf{f}$, llamada *hélice*, tiene el siguiente aspecto geométrico

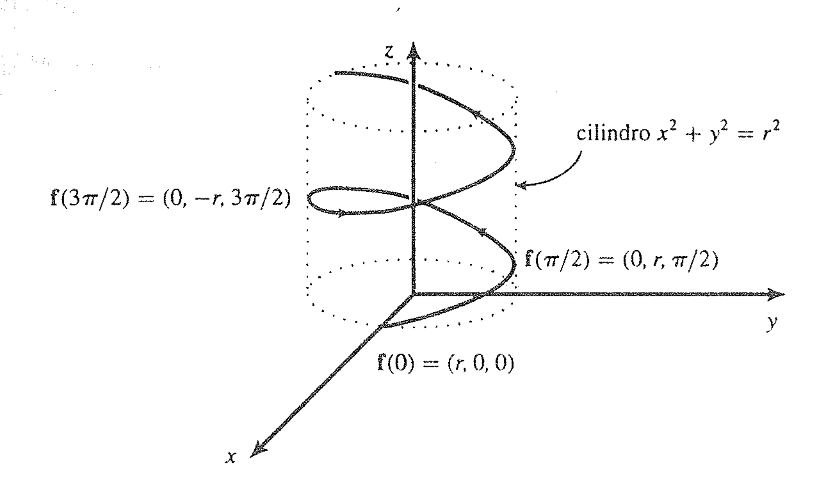


Figura 15. Una hélice. ■

Una terminología muy usada para las funciones coordenadas de caminos en $\mathbb{R}^2$ y $\mathbb{R}^3$, es llamar a éstas como las *ecuaciones paramétricas* de la curva correspondiente al camino. A la variable t se le llama parámetro. Por ejemplo, las ecuaciones $x(t) = r\cos t$, $y(t) = r\,\mathrm{sen}\,t$, $0 \le t \le 2\pi$ son las ecuaciones paramétricas de un círculo con centro en el origen y radio r (en el plano) —ver ejemplo 3; las ecuaciones $x(t) = a(t - \mathrm{sen}\,t)$, $y(t) = a(1 - \cos t)$, $t \in \mathbb{R}$, son las ecuaciones paramétricas de una cicloide —ver ejemplo 5, y las ecuaciones $x(t) = r\cos t$, $y(t) = r\,\mathrm{sen}\,t$, $z(t) = t$, $t \in \mathbb{R}$, son las ecuaciones paramétricas de una hélice.

Para adentrarnos en la exposición que emprenderemos a continuación, consideremos el ejemplo siguiente.

Sea $\bar{\mathbf{f}}: [0, \pi] \to \mathbb{R}^2$ el camino $\bar{\mathbf{f}}(t) = (r\cos 2t, r\,\mathrm{sen}\,2t)$. Se observa que las funciones coordenadas $x(t) = r\cos 2t$, $y(t) = r\,\mathrm{sen}\,2t$ satisfacen $x(t)^2 + y(t)^2 = r^2\ \forall t \in [0, \pi]$, que el punto inicial es $\bar{\mathbf{f}}(0) = (r, 0)$ y el punto final es $\bar{\mathbf{f}}(\pi) = (r, 0) = \bar{\mathbf{f}}(0)$. Así que la curva que representa este camino es un círculo con centro en el origen, radio r, que comienza su recorrido en el punto $(r, 0)$ y se recorre en sentido antihorario: exactamente la misma curva que representa el camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$,

$\mathbf{f}(t) = (r\cos t, r\,\mathrm{sen}\, t)$ estudiado en el ejemplo 3. Sin embargo, es claro que los caminos de $\mathbf{f}$ y $\bar{\mathbf{f}}$ son *distintos*: están definidos en distintos intervalos y las reglas de asociación de $t \in I$ con las imágenes correspondientes en $\mathbb{R}^2$ son distintas. Ambos, sin embargo, tienen la misma traza. La moraleja de este sencillo ejemplo introductorio es enfatizar que un camino no es solamente el conjunto de sus imágenes (i.e. la curva que representa). Habíamos ya dicho que podemos pensar en un camino como un objeto matemático caracterizado por dos aspectos: primero, la traza, curva, o forma geométrica del conjunto de sus imágenes, y segundo, *la manera como el camino recorre tal curva*. En el ejemplo anterior, vemos que tanto $\mathbf{f}$ como $\bar{\mathbf{f}}$ coinciden en el primer aspecto, pero no en el segundo: al ver los dominios de estas funciones y pensar en su variable independiente como tiempo, vemos que mientras $\mathbf{f}$ tarda 2π segundos en recorrer el círculo $x^2 + y^2 = r^2$, a $\bar{\mathbf{f}}$ sólo le toma la mitad de tiempo, π segundos, en efectuar tal recorrido. Esto sugiere que la velocidad de recorrido de $\bar{\mathbf{f}}$ es el doble de la de $\mathbf{f}$. El término "velocidad" aparece entonces de manera natural como el ingrediente que nos dice *cómo* una función $\mathbf{f}$ (un camino) recorre la curva de sus imágenes. Esto es precisamente lo que vamos a estudiar en la próxima sección.

Ejercicios (Capítulo 5, Sección 2)

1. Describa la traza del camino constante $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (a, b)$.

2. Sea $\varphi\colon I \subseteq \mathbb{R} \to \mathbb{R}$ una función continua definida en el intervalo I de $\mathbb{R}$. Considere el camino $\mathbf{f}\colon I \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, \varphi(t))$. Demuestre que éste debe ser un camino simple. ¿Puede ser un camino cerrado?

3. Demuestre que todo camino del tipo $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (at + b, ct + d)$, donde a y c son reales no nulos, es simple. Describa su traza.

4. Demuestre que el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t^2 + 1, t^2 - 1)$ no es simple. Describa la traza de $\mathbf{f}$.

5. Demuestre que el camino $\mathbf{g}\colon [1, \infty) \to \mathbb{R}^2$, $\mathbf{g}(t) = (t, t - 2)$ es simple. Describa la traza de $\mathbf{g}$ y compare con el ejercicio anterior.

6. Demuestre que el camino $\mathbf{f}\colon [-2, 2] \to \mathbb{R}^2$, $\mathbf{f}(t) = (t^2 - 1, t^3 + 2t^2 - t - 2)$ no es simple. ¿Es cerrado?, ¿ es cerrado simple?

7. Repita el ejercicio anterior con el camino $\mathbf{f}\colon [-1, 1] \to \mathbb{R}^2$, $\mathbf{f}(t) = (t^2 - 1, t^3 + 2t^2 - t - 2)$.

8. Considere el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (\cosh t, \mathrm{senh}\, t)$. Describa la traza de $\mathbf{f}$.

9. Considere el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (a, \varphi(t))$, donde a es un número real dado. Describa la traza de $\mathbf{f}$ en términos del rango de la función $\varphi\colon \mathbb{R} \to \mathbb{R}$. Suponga que ésta es una función inyectiva. ¿Qué puede decir del camino $\mathbf{f}$?, ¿qué puede decir de $\mathbf{f}$ si φ es sobreyectiva?, ¿si φ es biyectiva?

10. Repita el ejercicio anterior con el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (\varphi(t), a)$.

11. Sea $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$ un camino, y sea u un vector en $\mathbb{R}^2$ dado. Considere el camino $\mathbf{g}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{g}(t) = u + \mathbf{f}(t)$. Describa la traza de $\mathbf{g}$ en términos de la traza de $\mathbf{f}$.

12. Sean $\mathbf{u}_0$ y $\mathbf{u}_1$ dos vectores en $\mathbb{R}^2$. Considere el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = \mathbf{u}_0 + t\mathbf{u}_1$. Describa la traza de $\mathbf{f}$.

13. Sean $\mathbf{u}_0$, $\mathbf{u}_1$ y $\mathbf{u}_2$ tres vectores en $\mathbb{R}^2$. Considere el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = \mathbf{u}_0 + t\mathbf{u}_1 + t^2\mathbf{u}_2$. Describa la traza de $\mathbf{f}$ en cada uno de los casos: **a.** los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ son linealmente independientes, **b.** los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ son linealmente dependientes.

14. Describa la traza del camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = (a\cos t, b \operatorname{sen} t)$, en cada uno de los casos: **a.** $a = b$, **b.** $a \neq b$.

15. Sean $\mathbf{u}_0$, $\mathbf{u}_1$ y $\mathbf{u}_2$ tres vectores en $\mathbb{R}^2$. Considere el camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = \mathbf{u}_0 + \cos t\mathbf{u}_1 + \operatorname{sen} t\mathbf{u}_2$. Demuestre que si $\mathbf{u}_1$ y $\mathbf{u}_2$ son vectores linealmente independientes, la traza de $\mathbf{f}$ es una elipse. ¿Qué pasa si los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ son linealmente dependientes?

16. Considere el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t, \varphi(t), a)$, donde $\varphi: \mathbb{R} \to \mathbb{R}$ es una función continua. Describa la traza de $\mathbf{f}$.

17. Repita el ejercicio anterior con los caminos $\mathbf{g}, \mathbf{h}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{g}(t) = (t, a, \varphi(t))$, $\mathbf{h}(t) = (a, t, \varphi(t))$.

18. Describa la traza de los caminos $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}: [0, 2\pi] \to \mathbb{R}^3$, $\mathbf{f}(t) = (\cos t, \operatorname{sen} t, 0)$, $\mathbf{g}(t) = (\cos t, 0, \operatorname{sen} t)$, $\mathbf{h}(t) = (0, \operatorname{sen} t, \cos t)$.

19. Se dice que el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ se encuentra sobre la esfera $S = \{(x, y, z) | x^2 + y^2 + z^2 = R^2\}$ si $\mathbf{f}(t) \in S$ $\forall t \in I$ (es decir, si $f_1^2(t) + f_2^2(t) + f_3^2(t) = R^2$ $\forall t \in I$). Demuestre que los caminos $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$ del ejercicio anterior se encuentran sobre la esfera unitaria.

20. Se dice que la curva, imagen del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ es una *curva plana*, si hay un plano $S = \{(x, y, z) | ax + by + cz = d\}$ tal que $\mathbf{f}(t) \in S$ $\forall t \in I$ (en cuyo caso se dice que el camino $\mathbf{f}$ —o la curva que describe— se encuentra sobre el plano S). Considere los caminos $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$ de los ejercicios 16 y 17. Demuestre que las curvas que éstos describen son planas.

21. Demuestre que el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (a\cos^2 t, 0.5b \operatorname{sen}^2 t, 0.5c \operatorname{sen}^2 t)$ se encuentra sobre el plano $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$.

22. Demuestre que el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t^2 + t + 1, t^2 - 1, t + 2)$ se encuentra sobre el plano $z = x - y$.

23. El ejemplo anterior es un caso particular de la siguiente situación más general: la curva descrita por el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (a_1t^2 + b_1t + c_1, a_2t^2 + b_2t + c_2, a_3t^2 + b_3t + c_3)$ es una curva plana. Esta se encuentra sobre el plano cuya ecuación es

$$\det \begin{bmatrix} x - c_1 & y - c_2 & z - c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = 0$$

Demuestre este hecho.

5.3 Diferenciabilidad. Curvas Regulares.

El formalismo en la definición de la derivada de una función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ de las estudiadas en el primer curso de cálculo, dado por

$$f'(t) = \lim_{h \to 0} \frac{f(t + h) - f(t)}{h}$$

hace perfecto sentido si la función f, en lugar de tomar valores reales, toma valores en $\mathbb{R}^n$. Nada más natural entonces que definir, de esta misma manera, la derivada para caminos.

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino definido en el intervalo abierto I de $\mathbb{R}$. Sea $t_0 \in I$. Se define la derivada de $\mathbf{f}$ en t_0, denotada por $\mathbf{f}'(t_0)$ o $\frac{d\mathbf{f}}{dt}(t_0)$, como el límite

$$\mathbf{f}'(t_0) = \lim_{h \to 0} \frac{\mathbf{f}(t_0 + h) - \mathbf{f}(t_0)}{h}$$

cuando éste existe. En tal caso se dice que el camino $\mathbf{f}$ es diferenciable en t_0. Si $\mathbf{f}$ es diferenciable en todos los puntos t_0 de I, decimos que es diferenciable en I. ■

La primera observación que debemos hacer en este nuevo concepto, es que la derivada de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ en un punto $t_0 \in I$, *es un vector en el espacio* $\mathbb{R}^n$. De hecho, viendo la siguiente cadena de igualdades, en donde

$$f(t) = (x_1(t), x_2(t), \ldots, x_n(t))$$

$$\begin{aligned} f'(t_0) &= \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h} = \lim_{h \to 0} \frac{\big(x_1(t_0 + h), \ldots, x_n(t_0 + h)\big) - \big(x_1(t_0), \ldots, x_n(t_0)\big)}{h} \\ &= \lim_{h \to 0} \left(\frac{x_1(t_0 + h) - x_1(t_0)}{h}, \ldots, \frac{x_n(t_0 + h) - x_n(t_0)}{h} \right) \\ &= \left(\lim_{h \to 0} \frac{x_1(t_0 + h) - x_1(t_0)}{h}, \ldots, \lim_{h \to 0} \frac{x_n(t_0 + h) - x_n(t_0)}{h} \right) \\ &= \big(x_1'(t_0), \ldots, x_n'(t_0)\big) \end{aligned}$$

podemos concluir que el camino $\mathbf{f}$ es diferenciable en t_0 si y sólo si sus funciones coordenadas lo son, y, en tal caso, la derivada $\mathbf{f}'(t_0)$ es el vector de $\mathbb{R}^n$ cuyas coordenadas son las derivadas $x_i'(t_0)$ de las funciones coordenadas de $\mathbf{f}$.

NOTA. En el capítulo 3 se había dicho que la derivada de una función $\mathbf{f}: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ en un punto $\mathbf{x}_0$ del abierto U de $\mathbb{R}^n$ es una transformación lineal cuya representación en términos de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ es la matriz jacobiana de $\mathbf{f}$ en $\mathbf{x}_0$, la cual es una matriz de orden $m \times n$ que en su i-ésima línea y j-ésima columna tiene la derivada parcial de la i-ésima función coordenada de $\mathbf{f}$ respecto de su j-ésima variable. Con esta perspectiva, la derivada de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ en $t_0 \in I$, debería estar asociada a una matriz de orden $n \times 1$, la matriz jacobiana de $\mathbf{f}$ en t_0, en cuya i-ésima línea (tiene sólo una columna) se encuentra la derivada de la i-ésima función coordenada de $\mathbf{f}$ respecto de su (única) variable independiente t. Es decir, la derivada $\mathbf{f}'(t_0)$ del camino $\mathbf{f}$ en t_0, debe estar representada por su matriz jacobiana $J\mathbf{f}(t_0)$ la cual es

$$J\mathbf{f}(t_0) = \begin{bmatrix} x_1'(t_0) \\ x_2'(t_0) \\ \vdots \\ x_n'(t_0) \end{bmatrix}$$

Hemos establecido, de manera natural, el concepto de derivada del camino $\mathbf{f}$ en t_0 como el vector de $\mathbb{R}^n$ dado por

$$\mathbf{f}'(t_0) = \big(x_1'(t_0), x_2'(t_0), \ldots, x_n'(t_0)\big)$$

Por supuesto no hay ningún problema en conciliar estos dos puntos de vista para $\mathbf{f}'(t_0)$, identificando los vectores de $\mathbb{R}^n$ con matrices $n \times 1$, de la manera canónica. Esta identificación es de hecho un isomorfismo entre el espacio vectorial $\mathbb{R}^n$ y el espacio vectorial de las matrices de orden $n \times 1$.

Un hecho geométrico relevante acerca del vector $\mathbf{f}'(t_0)$ de un camino diferenciable $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$, es que éste es tangente a la curva correspondiente en el punto donde se calcula la derivada y que apunta en dirección al recorrido de la curva. Veamos, por ejemplo, el caso de un camino diferenciable en $\mathbb{R}^2$, $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$. En la siguiente figura están marcados los puntos de la curva correspondientes a t_0 y $t_0 + h$. El vector $\mathbf{v} = \mathbf{f}(t_0 + h) - \mathbf{f}(t_0)$ es aquél que va de P a Q. Para cualquier $h \neq 0$ el vector $\bar{\mathbf{v}} = \frac{1}{h}(\mathbf{f}(t_0 + h) - \mathbf{f}(t_0))$ es un vector paralelo a $\mathbf{v}$. Siguiendo la posición del vector $\bar{\mathbf{v}}$ para h cada vez más pequeño, resulta plausible que, en la posición límite, cuando $h \to 0$, el vector $\bar{\mathbf{v}}$ sea tangente a la curva en P.

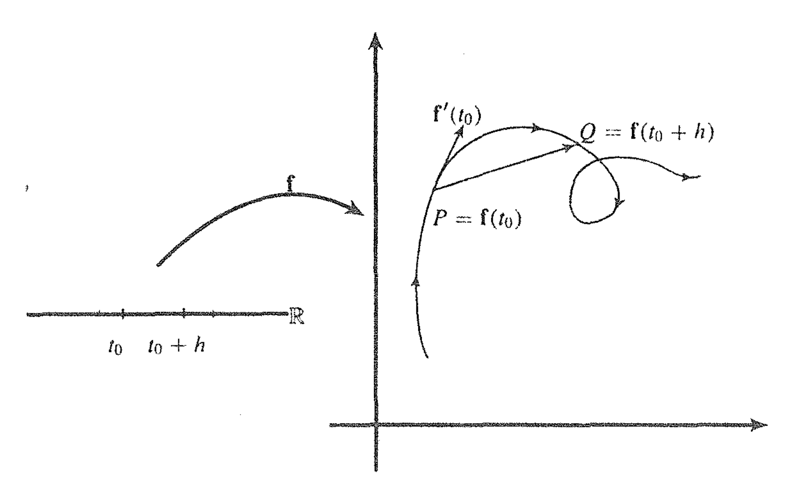


Figura 1. El vector tangente $\mathbf{f}'(t_0)$ a la curva en P.

Más aún, el vector tangente $\mathbf{f}'(t_0)$ mide la velocidad a que se trazan las imágenes en la curva, o bien, a qué velocidad se mueve un punto en el espacio siguiendo la trayectoria determinada por la función $\mathbf{f}$. La siguiente definición comprende este importante hecho.

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino diferenciable. Al vector $\mathbf{f}'(t)$ se le llama *vector velocidad* del camino en el punto $\mathbf{f}(t) \in \mathbb{R}^n$. ■

De esta manera, la diferenciabilidad de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ queda comprometida con el vector velocidad con el que la curva, traza del camino $\mathbf{f}$, es recorrida. Siendo éste tangente a la curva, usaremos indistintamente los nombres "vector velocidad" y "vector tangente".

Ejemplo 1. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (r\cos t, r\,\mathrm{sen}\, t)$. Este es un camino diferenciable,[1] pues sus funciones coordenadas $x = r\cos t$, $y = r\,\mathrm{sen}\, t$, lo son. La derivada de

[1] Se ha definido diferenciabilidad para caminos cuyos dominios son intervalos abiertos de $\mathbb{R}$. Cuando se consideran caminos definidos en intervalos cerrados, digamos $[a, b]$, podemos adecuar la definición de diferenciabilidad poniendo la función $\mathbf{f}$ definida en algún intervalo abierto un poco más grande que $[a, b]$, por ejemplo $(a - \epsilon, b + \epsilon) \supset [a, b]$ cuando la función misma lo permite (como en el ejemplo), o bien, se podrían definir las derivadas laterales $\mathbf{f}'_+(a)$ y $\mathbf{f}'_-(b)$ con los límites laterales correspondientes de la definición de derivada y decir entonces que el camino $\mathbf{f}$ es diferenciable en $[a, b]$ si lo es en (a, b) y existen las derivadas laterales $\mathbf{f}'_+(a)$ y $\mathbf{f}'_-(b)$.

f es

$$\mathbf{f}'(t) = (x'(t), y'(t)) = (-r \operatorname{sen} t, r \cos t)$$

Nótese que, tomando el producto punto del vector $\mathbf{f}(t)$ con el vector $\mathbf{f}'(t)$ se tiene

$$\begin{aligned} \mathbf{f}(t) \cdot \mathbf{f}'(t) &= (r \cos t, r \operatorname{sen} t) \cdot (-r \operatorname{sen} t, r \cos t) \\ &= -r^2 \operatorname{sen} t \cos t + r^2 \operatorname{sen} t \cos t = 0 \end{aligned}$$

de modo que $\mathbf{f}'(t)$ es perpendicular a $\mathbf{f}(t)$ para todo t. Puesto que la curva que describe f es el círculo $x^2 + y^2 = r^2$, resulta claro que $\mathbf{f}'(t)$ es tangente al círculo en el punto $\mathbf{f}(t)$ (el cual es un vector radial del círculo).

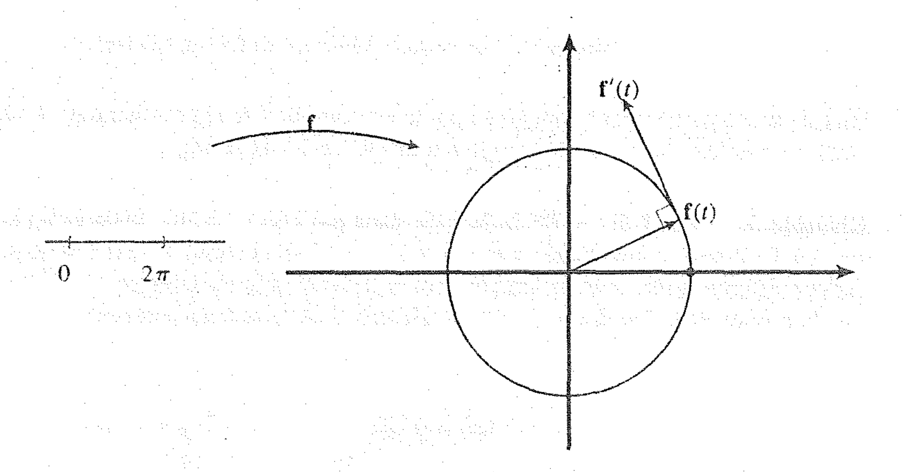


Figura 2. El camino $\mathbf{f}(t) = (r \cos t, r \operatorname{sen} t)$ y su vector velocidad $\mathbf{f}'(t)(-r \operatorname{sen} t, r \cos t)$. ■

Ejemplo 2. Consideremos el camino $\bar{\mathbf{f}}: [0, \pi] \to \mathbb{R}^2$ dado por $\bar{\mathbf{f}}(t) = (r \cos 2t, r \operatorname{sen} 2t)$. En la exposición con que cerramos la sección anterior veíamos que la curva descrita por $\bar{\mathbf{f}}$ es el mismo círculo $x^2 + y^2 = r^2$ descrito por el camino f del ejemplo anterior. Para este camino $\bar{\mathbf{f}}$ se tiene

$$\bar{\mathbf{f}}'(t) = (-2r \operatorname{sen} 2t, 2r \cos 2t)$$

En $t = 0$ ambos caminos se encuentran en el punto inicial $\mathbf{f}(0) = \bar{\mathbf{f}}(0) = (r, 0)$. Sin embargo, obsérvese que $\mathbf{f}'(0) = (0, r)$ y $\bar{\mathbf{f}}'(0) = (0, 2r) = 2\mathbf{f}'(0)$, de modo que la velocidad en el arranque de $\bar{\mathbf{f}}$ es el doble de la de $\mathbf{f}$. Esta es una situación que se mantiene durante todo el recorrido de las curvas correspondientes. En efecto, sea $\mathbf{p} = (a, b)$ un punto del círculo $x^2 + y^2 = r^2$, digamos $a = r \cos \theta$, $b = r \operatorname{sen} \theta$, con $0 \leq \theta \leq 2\pi$. El camino $\mathbf{f}$ alcanza al punto $\mathbf{p}$ a los $t = \theta$ seg. mientras que el camino $\bar{\mathbf{f}}$ llega a $\mathbf{p}$ a los $\bar{t} = \frac{1}{2}\theta$ seg. (pues $\mathbf{f}(\theta) = (a, b) = \bar{\mathbf{f}}(\frac{1}{2}\theta)$). Se tiene $\mathbf{f}'(\theta) = (-r \operatorname{sen} \theta, r \cos \theta)$ y $\bar{\mathbf{f}}'(\frac{1}{2}\theta) = (-2r \operatorname{sen} \theta, 2r \cos \theta) = 2\mathbf{f}'(\theta)$, es decir, al pasar por el punto $\mathbf{p}$ el punto que se mueve en el camino $\bar{\mathbf{f}}$ va al doble de la velocidad del que va por el camino $\mathbf{f}$.

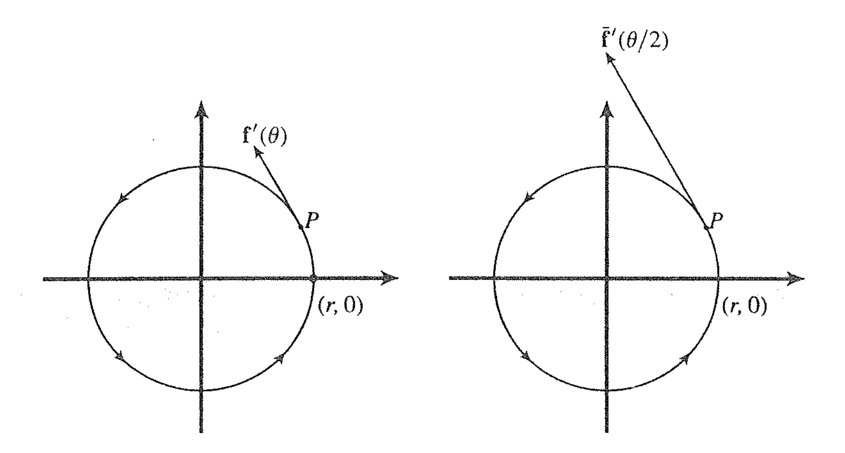


Figura 3. Los vectores velocidad de $\mathbf{f}$ y $\bar{\mathbf{f}}$ del ejemplo 2.

De esta manera podemos explicarnos por qué al camino $\bar{\mathbf{f}}$ le toma solamente π segundos recorrer el círculo completo, la mitad del tiempo del que le toma a $\mathbf{f}$(2π seg.). ■

Ejemplo 3. Sea $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$ la función dada por $\mathbf{f}(t) = (t, |t|)$. Este es un camino en $\mathbb{R}^2$ pues sus dos funciones coordenadas $x(t) = t$ y $y(t) = |t|$ son continuas. Sin embargo, no se trata de un camino diferenciable, pues la función $y(t) = |t|$ no es diferenciable en $t = 0$. La curva descrita por $\mathbf{f}$ es la gráfica de la función $y = |x|$ (ver ejemplo 1 de la sección anterior)

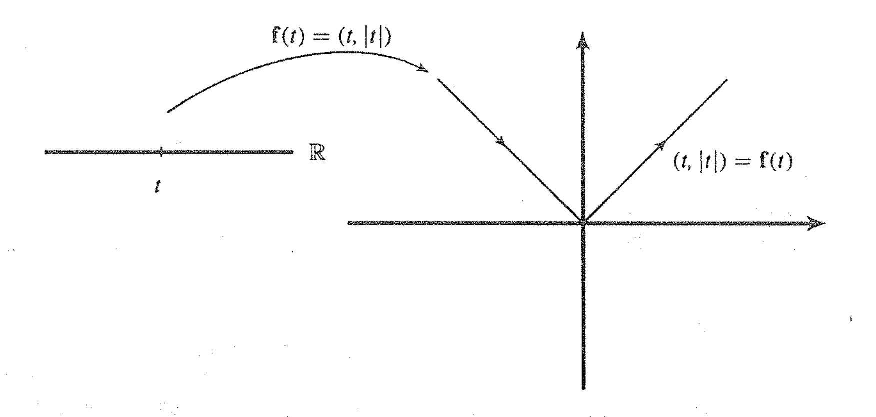


Figura 4. Traza de la función $\mathbf{f}(t) = (t, |t|)$.

Obsérvese que es justamente en el valor $t = 0$ en el que se presentan los problemas para el vector velocidad $\mathbf{f}'(t)$: en $\mathbf{f}(0) = (0, 0)$ dicho vector no existe. Sin embargo, para $t < 0$ se tiene $\mathbf{f}(t) = (t, -t)$ de modo que $\mathbf{f}'(t) = (1, -1)$ y para $t > 0$, se tiene $\mathbf{f}(t) = (t, t)$ de modo que $\mathbf{f}'(t) = (1, 1)$. Así pues el vector velocidad $\mathbf{f}'(t)$ está bien determinado, excepto para $t = 0$. ■

Ejemplo 4. Consideremos el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (t^3, t^2|t|)$. Ciertamente la función $x(t) = t^3$ es diferenciable; la función $y(t) = t^2|t|$ también lo es, pues para $t < 0$ se tiene $y(t) = -t^3$ que es diferenciable, para $t > 0$ se tiene $y(t) = t^3$ que también es diferenciable y para $t = 0$ tenemos

$$y'(0) = \lim_{h\to 0} \frac{y(h) - y(0)}{h} = \lim_{h\to 0} \frac{h^2|h| - 0}{h} = \lim_{h\to 0} h|h| = 0$$

Así que $y'(t)$ existe para todo $t \in \mathbb{R}$. Además observe cómo $|x(t)| = |t^3| = |t^2 t| = |t^2|\,|t| = t^2|t| = y(t)$, de modo que la curva descrita por este camino es la gráfica de la función $y = |x|$, misma del ejemplo anterior. Tenemos entonces que $\mathbf{f}$ es un camino diferenciable (sus funciones coordenadas lo son) y la curva a la que describe es la gráfica de la función $y = |x|$. Este ejemplo nos muestra que la diferenciabilidad de un camino nada tiene que ver con la forma geométrica de la curva que describe: la gráfica de $y = |x|$ es la curva descrita por el camino $\mathbf{g}(t) = (t, |t|)$ que no es diferenciable y es la misma curva descrita por el camino $\mathbf{f}(t) = (t^3, t^2|t|)$ que sí lo es.

Veamos el comportamiento del vector velocidad $\mathbf{f}'(t)$ para la curva de nuestro ejemplo. Para $t < 0$ tenemos $\mathbf{f}(t) = (t^3, -t^3)$, de donde $\mathbf{f}'(t) = (3t^2, -3t^2)$, para $t > 0$ tenemos $\mathbf{f}(t) = (t^3, t^3)$, de donde $\mathbf{f}'(t) = (3t^2, 3t^2)$ y para $t = 0$ tenemos $\mathbf{f}'(0) = (x'(0), y'(0)) = (0, 0)$, así que el vector velocidad $\mathbf{f}'(t)$ queda como

$$\mathbf{f}'(t) = \begin{cases} (3t^2, -3t^2) & \text{si } t < 0 \\ (0, 0) & \text{si } t = 0 \\ (3t^2, 3t^2) & \text{si } t > 0 \end{cases}$$

Imaginemos un punto $\mathbf{p}$ en $\mathbb{R}^2$ recorriendo el camino $\mathbf{f}(t) = (t^3, t^2|t|)$. Este viene desde el infinito por la recta $y = -x$ con una velocidad $\mathbf{f}'(t) = (3t^2, -3t^2)$. Al acercarse al origen la velocidad de $\mathbf{p}$ va disminuyendo (pues a medida que t es más pequeño, el vector $\mathbf{f}'(t)$ decrece en magnitud), de modo que al llegar a $(0, 0)$ el punto $\mathbf{p}$ *se detiene* para comenzar, para $t > 0$, su recorrido por la recta $y = x$ con una velocidad $\mathbf{f}'(t) = (3t^2, 3t^2)$ (la cual aumenta de magnitud conforme aumenta t). Este es entonces un recorrido diferenciable del punto $\mathbf{p}$ por la curva $y = |x|$. El hecho de que en el origen esta curva tenga un pico —situación que en el 1er. curso de cálculo asociábamos inmediatamente con la ausencia de diferenciabilidad— no afecta a la diferenciabilidad del recorrido. Lo único que ocurre

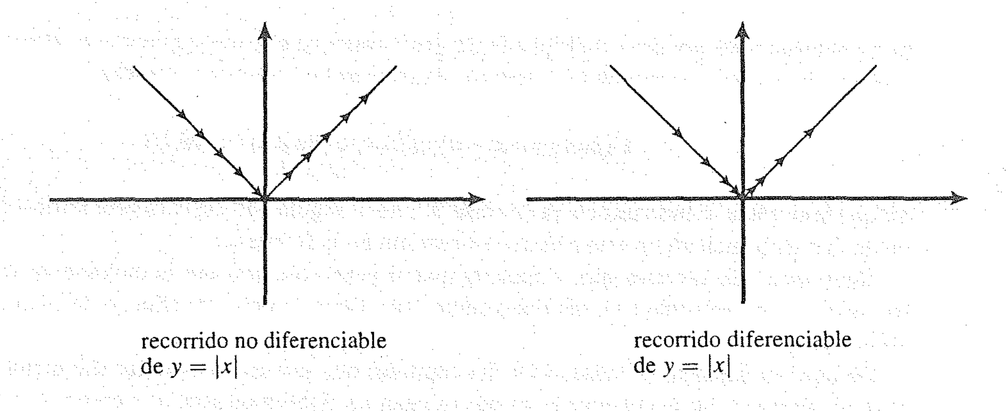


Figura 5. Los recorridos diferenciable y no diferenciable de $y = |x|$.

en ese pico, que es una conexión no suave entre las dos "carreteras" $y = -x$ y $y = x$, donde el punto que viene por la carretera $y = -x$, se detiene un instante (su velocidad es cero) para luego tomar la carretera $y = x$. Esta es la manera de "salvar diferenciablemente" los vértices en las carreteras: a medida que se llega a ellos se disminuye la velocidad, se hace una parada en el vértice y se arranca después en la nueva dirección. Esta situación no ocurre con el camino del ejemplo anterior, el cual representa un recorrido no diferenciable de la curva $y = |x|$: el punto viene por la carretera $y = -x$ con la velocidad constante dada por el vector $(1, -1)$ y luego cambia abruptamente a la carretera $y = x$ con la velocidad $(1, 1)$. ■

Los ejemplos anteriores muestran que la diferenciabiliad de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ no detecta picos de la curva que representa. Nótese que en el camino diferenciable del ejemplo 3, $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t^3, t^2|t|)$, al pasar por el pico de la curva $y = |x|$ (en $t = 0$) su velocidad se anula, es decir, en ese punto se tiene $\mathbf{f}'(0) = (0, 0)$. Por supuesto que este punto representa un problema geométrico al no poder asociar ahí una recta tangente a la curva. La propiedad de las curvas referente a la posibilidad de trazar rectas tangentes a ellas se llama "regularidad". Esta es una propiedad que definiremos para caminos de clase $\mathscr{C}^1$, los cuales, además de ser diferenciables, tienen continua su función derivada $\mathbf{f}': I \to \mathbb{R}^n$ (que a cada $t \in I$ asocia el vector tangente $\mathbf{f}'(t) \in \mathbb{R}^n$); (del mismo modo, las derivadas $x_1'(t)$ de las funciones coordenadas de $\mathbf{f}$ son continuas; es decir, las funciones $x_i(t)$ son de clase $\mathscr{C}^1$). La razón de hacerlo así es puramente técnica y no debe inquietar al lector. La mayor parte de los caminos con que trabajaremos no son sólo de clase $\mathscr{C}^1$, sino también de la clase $\mathscr{C}^\infty$ (con *todas* sus derivadas continuas).

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$. Se dice que éste es un *camino regular* si $\mathbf{f}'(t) \neq \mathbf{0}$ (el vector $\mathbf{0}$ de $\mathbb{R}^n$) para todo $t \in I$.

Así, el camino del ejemplo 3 es diferenciable (de hecho de clase $\mathscr{C}^1$) pero no es regular. Otro ejemplo interesante sobre este concepto es la cicloide (ver ejemplo 5 de la sección anterior), curva asociada con el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = a(t - \operatorname{sen} t, 1 - \cos t)$. No hay duda de que éste es un camino de la clase $\mathscr{C}^1$, pues las funciones coordenadas $x = a(t - \operatorname{sen} t)$, $y = a(1 - \cos t)$ lo son. El vector velocidad de esta curva es

$$\mathbf{f}'(t) = (x'(t), y'(t)) = (a - a\cos t, a \operatorname{sen} t)$$

Se ve entonces que por cada múltiplo de 2π (cada vez que el círculo que rueda sobre el eje x, cuyo punto en la periferia da origen a la cicloide, da una vuelta completa), se tiene

$$\mathbf{f}'(2k\pi) = (a - a\cos 2k\pi, a \operatorname{sen} 2k\pi) = (0, 0)$$

así que la cicloide, siendo una curva de clase $\mathscr{C}^1$, no es regular. El aspecto geométrico de los vectores velocidad en la cicloide es como el que se muestra en la figura 6.

Nuevamente se observa que, a medida que el punto que recorre la cicloide se va acercando a los vértices, su velocidad va disminuyendo, para llegar a cero en ellos y de nuevo aumentar al dejarlos atrás.

De aquí en adelante (y hasta el fin del capítulo) nos interesará estudiar solamente caminos que sean regulares. A las curvas que éstos representan les podremos asociar siempre *rectas tangentes* y *planos normales* (para caminos en $\mathbb{R}^3$) que estudiaremos a continuación.

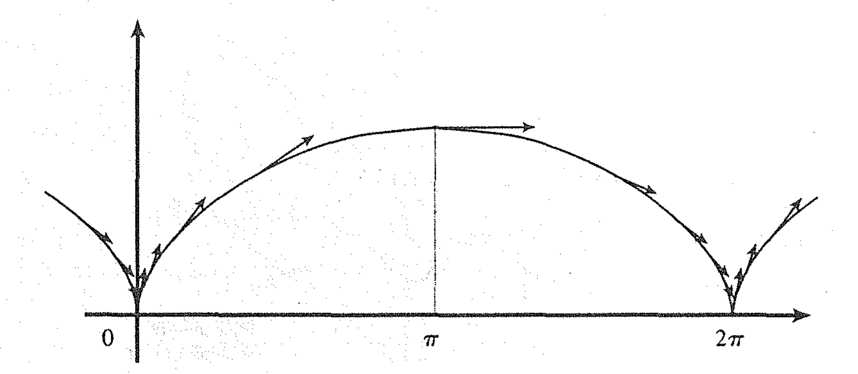


Figura 6. Los vectores velocidad en la cicloide.

Sea $\mathbf{f}: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$ un camino regular. Sea $t_0 \in I$. La *recta tangente* a la curva correspondiente en $\mathbf{f}(t_0)$ es la recta en $\mathbb{R}^3$ que pasa por $\mathbf{f}(t_0)$ y tiene por vector paralelo a $\mathbf{f}'(t_0)$. Entonces, si $\mathbf{f}(t) = (x(t), y(t), z(t))$, tenemos que la recta tangente a la curva en $\mathbf{f}(t_0)$ es

$$\begin{cases} x = x(t_0) + x'(t_0)s \\ y = y(t_0) + y'(t_0)s, \\ z = z(t_0) + z'(t_0)s \end{cases} \qquad s \in \mathbb{R}$$

(Obsérvese que también podemos decir que tal recta es la imagen del camino $\mathbf{g}: \mathbb{R} \rightarrow \mathbb{R}^3$, $\mathbf{g}(s) = (x(t_0)+x'(t_0)s, y(t_0)+y'(t_0)s, z(t_0)+z'(t_0)s) = \mathbf{f}(t_0)+\mathbf{f}'(t_0)s$). Quitando la última coordenada tenemos la correspondiente recta tangente a una curva en el plano (figura 7).

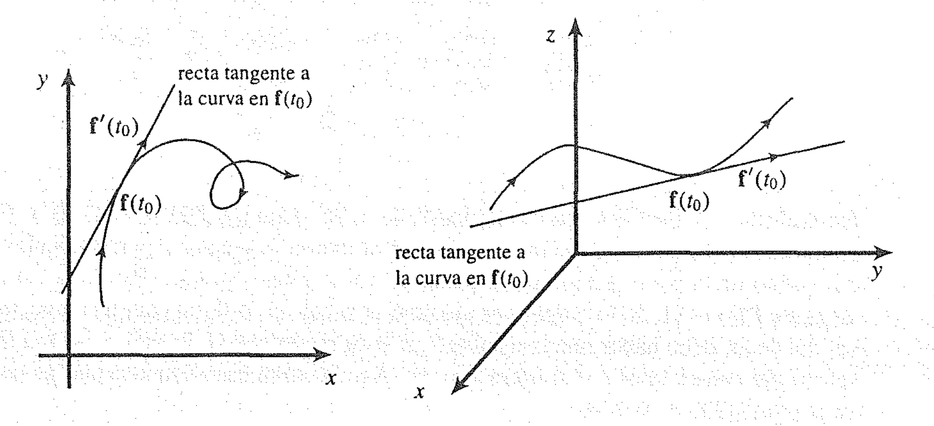


Figura 7. La recta tangente a una curva.

Para el camino $\mathbf{f}: I \subseteq \mathbb{R} \rightarrow \mathbb{R}^3$, el *plano normal* a la curva correspondiente en $\mathbf{f}(t_0)$ es el plano en $\mathbb{R}^3$ que pasa por $\mathbf{f}(t_0)$ y tiene ahí por vector normal al vector $\mathbf{f}'(t_0)$. Así la ecuación de tal plano (figura 8) es

$$x'(t_0)(x - x(t_0)) + y'(t_0)(y - y(t_0)) + z'(t_0)(z - z(t_0)) = 0$$

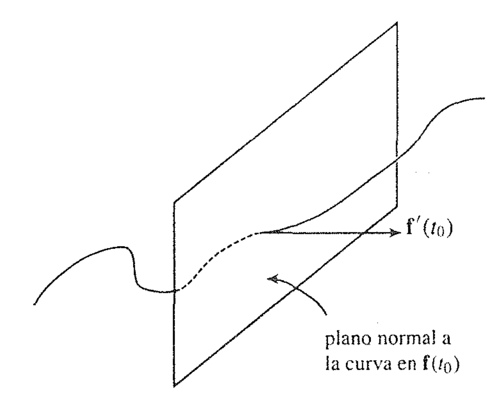


Figura 8. El plano normal a una curva.

Ejemplo 5. Consideremos la hélice $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (\cos t, \operatorname{sen} t, t)$. Hallemos la recta tangente y el plano normal a esta curva en el punto $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{\pi}{4}\right) = \mathbf{f}\left(\frac{\pi}{4}\right)$. Tenemos que el vector velocidad $\mathbf{f}'(t) = (-\operatorname{sen} t, \cos t, 1)$ de modo que en $t = \frac{\pi}{4}$ se tiene $\mathbf{f}'\left(\frac{\pi}{4}\right) = \left(\frac{-\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1\right)$. Así, la recta tangente procurada es

$$\begin{cases} x = \dfrac{\sqrt{2}}{2} - \dfrac{\sqrt{2}}{2}s \\ y = \dfrac{\sqrt{2}}{2} + \dfrac{\sqrt{2}}{2}s \\ z = \dfrac{\pi}{4} + s \end{cases}, \qquad s \in \mathbb{R}$$

y el plano normal

$$-\frac{\sqrt{2}}{2}\left(x - \frac{\sqrt{2}}{2}\right) + \frac{\sqrt{2}}{2}\left(y - \frac{\sqrt{2}}{2}\right) + \left(z - \frac{\pi}{4}\right) = 0$$

$$z = \frac{\pi}{4} + \frac{\sqrt{2}}{2}(x - y)$$

■

Ejemplo 6. Consideremos el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por $\mathbf{f}(t) = (t, t^2, 2t^3)$. Se quiere saber en qué punto (si hay) la curva correspondiente atraviesa el plano yz perpendicularmente, o bien, en qué punto de la curva ésta tiene al plano yz como plano normal. Para que esto ocurra, el vector tangente $\mathbf{f}'(t) = (1, 2t, 6t^2)$ debe ser paralelo al vector $(1, 0, 0)$, el cual es un vector normal al plano yz. Es decir, debe haber una constante $k \neq 0$ de modo que $(1, 2t, 6t^2) = k(1, 0, 0)$. Esta condición se cumple con el valor $t = 0$ (para $k = 1$). Así, la curva atraviesa el plano yz perpendicularmente en el punto $\mathbf{f}(0) = (0, 0, 0)$. ■

La propiedad de que una recta dada (en $\mathbb{R}^3$ o en $\mathbb{R}^2$) sea tangente a una curva, debe ser una propiedad de la curva en sí y no del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ (o $\mathbb{R}^3$) que tiene por traza dicha curva. Este es un hecho que expondremos ampliamente cuando estudiemos las reparametrizaciones de caminos en la sección 4. Esto mismo se puede decir para un plano normal a una curva en el espacio $\mathbb{R}^3$. Los dos siguientes ejemplos tienen que ver con este hecho.

Ejemplo 7. Consideremos los caminos $\mathbf{f}_n\colon \left[0, \frac{2\pi}{n}\right] \to \mathbb{R}^2$ dados por

$$\mathbf{f}_n(t) = (\cos nt, \operatorname{sen} nt) \qquad (n \in \mathbb{N})$$

Todos inician su recorrido en el punto $\mathbf{f}(0) = (1, 0)$ y lo terminan en $\mathbf{f}\left(0, \frac{2\pi}{n}\right) = (1, 0) = \mathbf{f}(0)$, siguiendo en sentido antihorario el círculo $x^2 + y^2 = 1$. Tomemos un punto (a, b) sobre dicho círculo, digamos $a = \cos\theta$, $b = \operatorname{sen}\theta$. El camino $\mathbf{f}_n$ pasa por (a, b) cuando $t = \frac{\theta}{n}$. El vector velocidad en este punto es $\mathbf{f}'\left(\frac{\theta}{n}\right) = \left(-n\operatorname{sen}(\frac{\theta}{n}n), n\cos(\frac{\theta}{n}n)\right) = (-n\operatorname{sen}\theta, n\cos\theta) = n(-\operatorname{sen}\theta, \cos\theta)$. Se ve entonces que "esencialmente" el vector tangente a la circunferencia $x^2 + y^2 = 1$ en (a, b) es el vector $(-\operatorname{sen}\theta, \cos\theta)$ pues todos los demás son múltiplos de él. La propiedad de tangencia de la recta que pasa por (a, b) y tiene a $\mathbf{f}'\left(\frac{\theta}{n}\right)$ como vector paralelo no dependerá entonces del valor de n (figura 9).

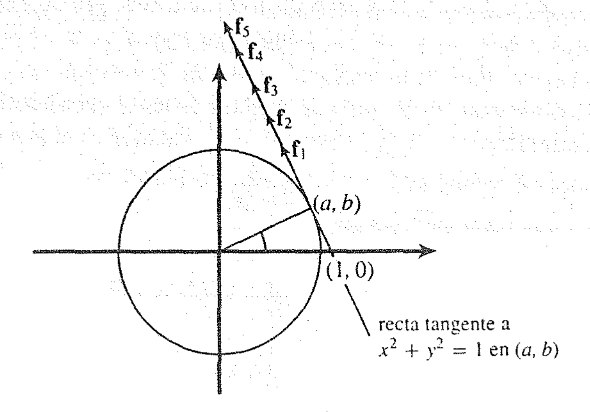


Figura 9. El vector tangente a $x^2 + y^2 = 1$ en (a, b). ■

Ejemplo 8. Sea C la curva que se obtiene de la intersección de la esfera $x^2 + y^2 + z^2 = 1$ y el plano $z = y$. Entonces C es una circunferencia en el espacio. Se quiere obtener la ecuación de la recta tangente y el plano normal a C en el punto $\mathbf{p} = \left(\frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2}\right)$ (figura 10).

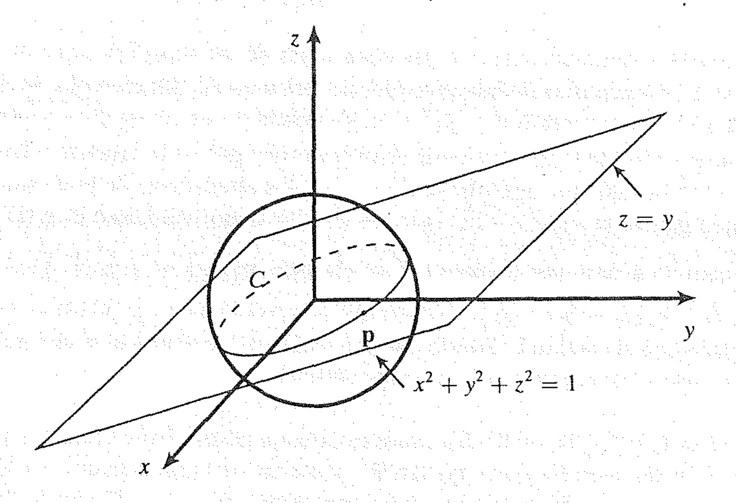


Figura 10. Intersección de la esfera y el plano.

Las ecuaciones de la recta tangente y el plano normal a una curva C en el espacio están planteadas en términos de las funciones coordenadas de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ cuya traza sea la curva C, así como de sus derivadas. En este ejemplo se nos da la descripción geométrica de la curva C pero no se nos da un camino que la tenga por traza. Sin embargo, como ya señalábamos, lo que nuestro ejemplo pide no deberá depender del camino concreto que tomemos que describa la curva C en los alrededores de $\mathbf{p}$. Todo lo que tenemos que hacer es determinar *un* camino (cualquiera) $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ que describa la curva C en los alrededores de $\mathbf{p}$. Esto lo podemos hacer de muchas maneras. La idea general es resolver simultáneamente las ecuaciones $x^2 + y^2 + z^2 = 1$, $z = y$ que definen a la curva y declarar alguna de las variables como t —el argumento de la función $\mathbf{f}$ que queremos hallar. Por ejemplo, poniendo $z = y = t$, nos queda $x^2 + t^2 + t^2 = 1$, o sea $x^2 = 1 - 2t^2$. En este momento surge el problema de escribir $x = \sqrt{1 - 2t^2}$ o $x = -\sqrt{1 - 2t^2}$. Al observar que el punto $\mathbf{p}$ donde queremos obtener la ecuación de la recta tangente y el plano normal tiene su primera coordenada positiva, nos decidimos por $x = \sqrt{1 - 2t^2}$. Así, el camino $\mathbf{f}: I \to \mathbb{R}^3$ definido en alguna vecindad I de $(y = z =)\ t = \frac{1}{2}$ dado por $\mathbf{f}(t) = \left(\sqrt{1 - 2t^2}, t, t\right)$ describe la traza de la curva que nos interesa. Este es un camino regular en I (tomado suficientemente pequeño). Nótese que no lo es globalmente en $\mathbb{R}$, pues su primera función coordenada $x(t) = \sqrt{1 - 2t^2}$tiene problemas de derivabilidad en $t = \pm\frac{1}{\sqrt{2}}$. Cuando $t = \frac{1}{2}$ estamos en el punto $\mathbf{p}$, es decir, $\mathbf{f}'\left(\frac{1}{2}\right) = \left(\frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2}\right) = \mathbf{p}$. El vector velocidad es $\mathbf{f}'(t) = \left(\frac{-2t}{\sqrt{1-2t^2}}, 1, 1\right)$ que para $t = \frac{1}{2}$ se ve como $\mathbf{f}'\left(\frac{1}{2}\right) = (-\sqrt{2}, 1, 1)$, así que la recta tangente buscada es

$$\begin{cases} x = \frac{1}{\sqrt{2}} - \sqrt{2}t \\ y = \frac{1}{2} + t, \\ z = \frac{1}{2} + t \end{cases} \qquad t \in \mathbb{R}$$

y el plano normal es

$$-\sqrt{2}\left(x - \frac{1}{\sqrt{2}}\right) + \left(y - \frac{1}{2}\right) + \left(z - \frac{1}{2}\right) = 0$$

o sea

$$-\sqrt{2}x + y + z = 0$$

Al resolver simultáneamemnte las ecuaciones de las superficies que al intersectarse determinan la curva C, hubiéramos podido proceder de otra manera, por ejemplo, declarando $x = t$. En tal caso, como $y = z$, se tendría $t^2 + 2y^2 = 1$, de donde $y = z = \frac{1}{\sqrt{2}}\sqrt{1 - t^2}$ (nuevamente optamos por el signo positivo de la raíz cuadrada atendiendo al signo de la segunda y tercera coordenadas del punto $\mathbf{p}$). Otro camino que describe la curva C en los alrededores de $\mathbf{p}$ es entonces el camino $\mathbf{g}: J \to \mathbb{R}^3$, donde $\mathbf{g}(t) = \left(t, \frac{1}{\sqrt{2}}\sqrt{1 - t^2}, \frac{1}{\sqrt{2}}\sqrt{1 - t^2}\right)$. Su vector velocidad es $\mathbf{g}'(t) = \left(1, \frac{-t}{\sqrt{2}\sqrt{1-t^2}}, \frac{-t}{\sqrt{2}\sqrt{1-t^2}}\right)$. El camino $\mathbf{g}$ pasa por $\mathbf{p}$ cuando $t = \frac{1}{\sqrt{2}}$ pues $\mathbf{g}\left(\frac{1}{\sqrt{2}}\right) = \left(\frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2}\right) = \mathbf{p}$. En ese punto se tiene $\mathbf{g}'\left(\frac{1}{\sqrt{2}}\right) = \left(1, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$. Nótese que este vector es $-\frac{1}{\sqrt{2}}$ veces el vector velocidad $(-\sqrt{2}, 1, 1)$ hallado con el camino $\mathbf{f}$. Puesto que son paralelos, se tendrá la misma recta tangente y plano normal a la curva C en el punto $\mathbf{p}$, como esperábamos. ■

Sean $\mathbf{f}, \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ dos caminos diferenciables. Para cada $t \in I$ podemos formar el producto punto de los vectores $\mathbf{f}(t)$ y $\mathbf{g}(t)$ en $\mathbb{R}^n$, y definir así la función $\varphi: I \to \mathbb{R}$ dada por $\varphi(t) = \mathbf{f}(t) \cdot \mathbf{g}(t)$. Si $x_i: I \to \mathbb{R}$ son las funciones coordenadas de $\mathbf{f}$ y $y_i: I \to \mathbb{R}$ son las funciones coordenadas de $\mathbf{g}$,

$i = 1, 2, \ldots, n$, entonces

$$\varphi(t) = \mathbf{f}(t) \cdot \mathbf{g}(t) = \sum_{i=1}^{n} x_i(t) y_i(t)$$

Ciertamente la función φ es una función diferenciable y su derivada es

$$\begin{aligned}
\varphi'(t) &= \frac{d}{dt} \sum_{i=1}^{n} x_i(t) y_i(t) = \sum_{i=1}^{n} \frac{d}{dt}\big(x_i(t) y_i(t)\big) \\
&= \sum_{i=1}^{n} \big(x_i(t) y_i'(t) + y_i x_i'(t)\big) \\
&= \sum_{i=1}^{n} x_i(t) y_i'(t) + \sum_{i=1}^{n} y_i(t) x_i'(t) \\
&= \mathbf{f}(t) \cdot \mathbf{g}'(t) + \mathbf{g}(t) \cdot \mathbf{f}'(t)
\end{aligned}$$

es decir que

$$\frac{d}{dt}\big(\mathbf{f}(t) \cdot \mathbf{g}(t)\big) = \mathbf{f}(t) \cdot \mathbf{g}'(t) + \mathbf{g}(t) \cdot \mathbf{f}'(t)$$

obteniéndose así una sugestiva fórmula para la derivada del producto punto de dos caminos diferenciables, que nos recuerda la fórmula de la derivada del producto de dos funciones reales de una variable real.

Como caso particular, poniendo $\mathbf{g}(t) = \mathbf{f}(t)$, obtenemos

$$\frac{d}{dt}\big(\mathbf{f}(t) \cdot \mathbf{f}(t)\big) = \mathbf{f}(t)\mathbf{f}'(t) + \mathbf{f}(t) \cdot \mathbf{f}'(t) = 2\mathbf{f}(t)\mathbf{f}'(t)$$

es decir

$$\frac{d}{dt}\|\mathbf{f}(t)\|^2 = 2\mathbf{f}(t) \cdot \mathbf{f}'(t)$$

Por otra parte, usando la regla de la cadena con la función $\varphi(t) = \|\mathbf{f}(t)\|^2$, vista como la composición de la función $\alpha(t) = t^2$ con $\beta(t) = \|\mathbf{f}(t)\|$ tenemos

$$\frac{d}{dt}\|\mathbf{f}(t)\|^2 = 2\|\mathbf{f}(t)\| \frac{d}{dt}\|\mathbf{f}(t)\|$$

y entonces

$$2\|\mathbf{f}(t)\| \frac{d}{dt}\|\mathbf{f}(t)\| = 2\mathbf{f}(t) \cdot \mathbf{f}'(t)$$

de donde, si $\mathbf{f}(t) \neq \mathbf{0}$, $t \in I$ (nos referimos a que $\mathbf{f}(t)$ no es el vector $\mathbf{0}$ de $\mathbb{R}^n$ para $t \in I$, lo cual equivale a que $\|\mathbf{f}(t)\| \neq 0$; esto significa que la curva que representa el camino $\mathbf{f}$ no pasa por el origen), obtenemos la fórmula de la derivada de la función $\psi(t) = \|\mathbf{f}(t)\|$ como

$$\frac{d}{dt}\|\mathbf{f}(t)\| = \frac{\mathbf{f}(t) \cdot \mathbf{f}'(t)}{\|\mathbf{f}(t)\|}$$

Ejemplo 9. En geometría analítica se define una circunferencia con centro en el origen como el lugar geométrico de todos los puntos del plano que equidistan del origen. Esta distancia común es

el radio r de la circunferencia. Podríamos definir la circunferencia, a la luz de lo estudiado en este capítulo, como la curva cerrada, imagen del camino regular $\mathbf{f}: I \to \mathbb{R}^2$ de modo que $\|\mathbf{f}(t)\| = r$, $\forall t \in I$. Esquemáticamente (figura 11)

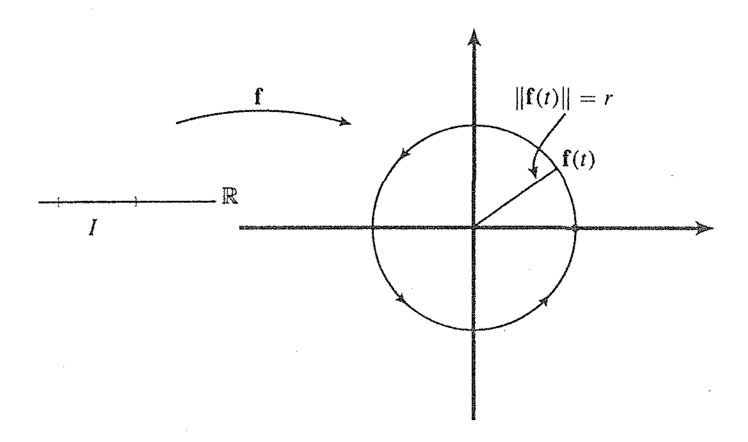


Figura 11. Una circunferencia.

Al derivar ambos miembros de la ecuación $\|\mathbf{f}(t)\| = r$, obtenemos, según la fórmula establecida anteriormente

$$\frac{\mathbf{f}(t) \cdot \mathbf{f}'(t)}{\|\mathbf{f}(t)\|} = 0 \qquad \forall t \in I$$

de donde $\mathbf{f}(t) \cdot \mathbf{f}'(t) = 0 \; \forall t \in I$. Concluimos entonces que los vectores $\mathbf{f}(t)$ y $\mathbf{f}'(t)$ son ortogonales en todos los puntos de la curva. Este es un hecho geométrico que ya conocíamos para una circunferencia

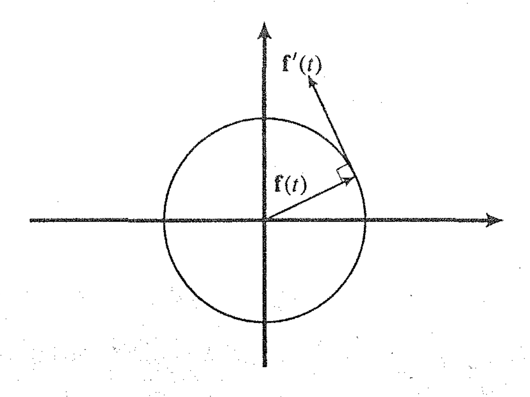


Figura 12. Los vectores $\mathbf{f}(t)$ y $\mathbf{f}'(t)$ en una circunferencia.

Más aún, supongamos que el camino regular $\mathbf{f}: I \to \mathbb{R}^2$ describe una curva cerrada definida en el intervalo I de $\mathbb{R}$, y es tal que $\mathbf{f}(t) \cdot \mathbf{f}'(t) = 0 \; \forall t \in I$. Entonces, por la misma fórmula anterior tenemos que la función diferenciable $\psi(t) = \|\mathbf{f}(t)\|$, definida en el intervalo I, tiene derivada igual a cero para toda $t \in I$. Usando el teorema del valor medio (para funciones reales de una variable

real) concluimos que $\psi(t) = \|\mathbf{f}(t)\| = \text{cte. } t \in I$. Es decir, la curva que describe el camino **f** es una circunferencia con centro en el origen. Así pues, podríamos definir una circunferencia como la imagen del camino cerrado regular $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ para el que $\mathbf{f}(t) \cdot \mathbf{f}'(t) = 0 \; \forall t \in I$. ■

Ejemplo 10. Sea $\mathbf{f}: I \to \mathbb{R}^2$ un camino cerrado regular en el plano. De manera intuitiva resulta claro que los vectores velocidad $\mathbf{f}'(t)$ toman todas las direcciones posibles.

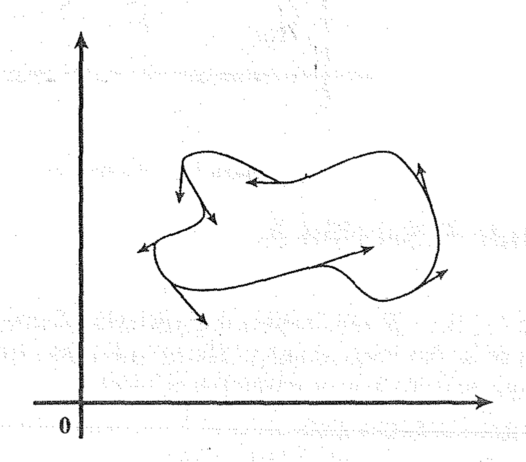


Figura 13. Los vectores velocidad en una curva cerrada en $\mathbb{R}^2$.

Este es un hecho que se puede probar rigurosamente. En efecto, sea $\mathbf{v} = (x_0, y_0)$ un vector no nulo arbitrario (pero fijo) en $\mathbb{R}^2$. Sea $\bar{\mathbf{v}} = (-y_0, x_0)$. Este es un vector ortogonal a **v**. Considere la función $\mathbf{g}: I \to \mathbb{R}^2$ dada por $\mathbf{g}(t) = \bar{\mathbf{v}} \cdot \mathbf{f}(t)$. Siendo **f** un camino cerrado, la función **g** toma el mismo valor en los extremos del intervalo $I = [a, b]$ (pues $\mathbf{g}(a) = \bar{\mathbf{v}} \cdot \mathbf{f}(a) = \bar{\mathbf{v}} \cdot \mathbf{f}(b) = \mathbf{g}(b)$). Se verifica fácilmente que las demás hipótesis del Teorema de Rolle se cumplen también para la función **g**. Según este teorema, debe existir $t_0 \in I$ de modo que $\mathbf{g}'(t_0) = 0$, es decir, tal que $\bar{\mathbf{v}} \cdot \mathbf{f}'(t_0) = 0$. En tal punto el vector $\mathbf{f}'(t_0)$ es perpendicular a $\bar{\mathbf{v}}$ y por lo tanto paralelo a **v**. Como **v** fue arbitrario, con esto concluimos que $\mathbf{f}'(t)$ toma todas las direcciones posibles, como queríamos. ■

Ejemplo 11. Sea $\mathbf{f}: I \to \mathbb{R}^3$ un camino regular de modo que $\mathbf{f}(t) \neq 0 \; \forall t \in I$. Supongamos que hay un $t_0 \in I$ para el que la distancia del punto $\mathbf{f}(t_0) \in \mathbb{R}^3$ al origen alcanza un valor mínimo. Afirmamos que en ese punto el vector $\mathbf{f}'(t_0)$ es perpendicular a $\mathbf{f}(t_0)$. En efecto, si consideramos la función $\psi(t) = \|\mathbf{f}(t)\| =$ distancia del punto $\mathbf{f}(t_0)$ al origen, puesto que ésta alcanza un mínimo en t_0 se tiene que $\psi'(t_0) = 0$. Entonces, según la fórmula de la derivada de $\|\mathbf{f}(t)\|$ tenemos

$$\psi'(t_0) = \frac{\mathbf{f}(t_0) \cdot \mathbf{f}'(t_0)}{\|\mathbf{f}(t_0)\|} = 0$$

de donde $\mathbf{f}(t_0) \cdot \mathbf{f}'(t_0) = 0$, como queríamos ver. Este es un hecho igualmente válido para curvas en el plano. Esquemáticamente (figura 14)

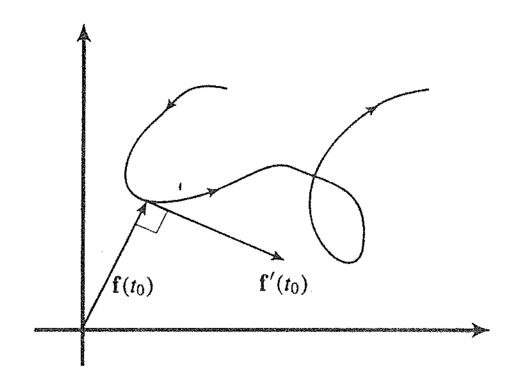


Figura 14. Ejemplo 11. ■

Ejercicios (Capítulo 5, Sección 3)

1. Sean $\mathbf{f}, \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ dos caminos diferenciables. Demuestre las siguientes fórmulas para las derivadas de las funciones suma y producto cruz de los caminos **f** y **g** (la derivada de la función producto punto de **f** y **g** ya se efectuó en el texto).

a. $(\mathbf{f} + \mathbf{g})'(t) = \mathbf{f}'(t) + \mathbf{g}'(t)$

b. $(\mathbf{f} \times \mathbf{g})'(t) = \mathbf{f}(t) \times \mathbf{g}'(t) + \mathbf{f}'(t) \times \mathbf{g}(t)$

En los ejercicios 2–5, hallar la derivada del camino dado en el punto indicado

2. $\mathbf{f}(t) = (\operatorname{sen}^2 t, \operatorname{sen} t^2)$, para $t = \pi/2$

3. $\mathbf{f}(t) = (\sqrt{3t+1}, \ln^3(t^2+1))$, para $t = 0$

4. $\mathbf{f}(t) = (\cos t, \cos^2 2t, \cos^3 3t)$, para $t = \pi$

5. $\mathbf{f}(t) = (\exp(-0.5t^2), -0.5t^2, \ln(t^2 + 0.5))$, para $t = 1$

En los ejercicios 6–10, diga si el camino es: **a.** diferenciable, **b.** regular. Justifique su respuesta.

6. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (3t - 1, 4t + 5)$

7. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (2t^2 + 4, 4t^3 - 2t^2 + 1)$

8. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^5 + 3t^4 - 5t^3 + 9t^2 - 10t + 23)$

9. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^{2/3})$

10. $f: \mathbb{R} \to \mathbb{R}$, $\mathbf{f}(t) = (t^3, t^2)$

11. ¿Cierto o falso? Si $\mathbf{f}, \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ son dos caminos regulares, entonces $\mathbf{f} + \mathbf{g}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es un camino regular.

12. ¿Cierto o falso? Si $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es un camino diferenciable tal que $\mathbf{f}(t) \neq 0\ \forall t \in I$, entonces la función $\varphi: I \subseteq \mathbb{R} \to \mathbb{R}$, $\varphi(t) = \|\mathbf{f}(t)\|$ es diferenciable y $\|\mathbf{f}'(t)\| = \|\mathbf{f}(t)\|'$ (es decir, $\varphi(\mathbf{f}'(t)) = \varphi'(\mathbf{f}(t))$). Si es cierto, demuéstrelo. Caso contrario, dé un contraejemplo.

13. ¿Cierto o falso? Si $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es un camino diferenciable tal que $\mathbf{f}(t) \neq 0\ \forall t \in I$, entonces $\mathbf{f}(t) \cdot \mathbf{f}'(t) = \|\mathbf{f}(t)\| \|\mathbf{f}'(t)\|$. Si es cierto, demuéstrelo. Caso contrario, dé un contraejemplo.

14. ¿Cierto o falso? Si $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es un camino diferenciable tal que $\mathbf{f}(t) \neq 0\ \forall t \in I$, entonces $\mathbf{f}(t) \cdot \mathbf{f}'(t) = \|\mathbf{f}(t)\| \|\mathbf{f}(t)\|'$. Si es cierto, demuéstrelo. Caso contrario, dé un contraejemplo.

15. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (a \cos t, b \operatorname{sen} t)$, donde a y b son dos reales positivos. Demuestre que $\mathbf{f}(t)$ es perpendicular a $\mathbf{f}'(t)$ para todo t en $[0, 2\pi]$ si y sólo si $a = b$. Interprete este hecho geométricamente (ver ejemplo 9).

16. Hallar el (los) valor(es) de t para los cuales el vector tangente al camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (2t^2 + 1, 3t - 2)$, sea paralelo al vector $\mathbf{v} = (2, -1)$.

17. Considere el camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = (a \cos t, b \operatorname{sen} t)$, en donde a y b son dos números reales positivos. Compruebe que se trata de un camino regular cerrado (¿cuál es su traza?). En el ejemplo 10 se demostró que el vector tangente $\mathbf{f}'(t)$ toma todas las direcciones posibles. Determine el (los) valores de $t \in [0, 2\pi]$ en los que $\mathbf{f}'(t)$ toma las direcciones de los siguientes vectores: **a.** $\mathbf{i}$; **b.** $-\mathbf{i}$; **c.** $\mathbf{j}$; **d.** $-\mathbf{j}$; **e.** $(1, 1)$; **f.** $(-1, -1)$.

18. Considerando el camino regular $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, \varphi(t))$, demuestre que la ecuación de la recta tangente a la gráfica de la función $\varphi: I \subseteq \mathbb{R} \to \mathbb{R}$ en el punto $x = x_0 \in I$, es

$$y = \varphi'(x_0)(x - x_0) + \varphi(x_0)$$

En los ejercicios 19–25, determine la ecuación de la recta tangente y del plano normal a la curva descrita por el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$, en el punto indicado

19. $\mathbf{f}(t) = (t + 1, 3t - 1, 2t - 5)$, en el punto $\mathbf{p} = \mathbf{f}(t_0)$

20. $\mathbf{f}(t) = (t^3 - 2t, t^2 + 1, 3)$, en el punto $\mathbf{p} = \mathbf{f}(1)$

21. $\mathbf{f}(t) = (\cos 3t, \operatorname{sen} 3t, 3t)$, en el punto $\mathbf{p} = \mathbf{f}(0)$

22. $\mathbf{f}(t) = (e^{-2t}, e^{-t}, t)$, en el punto $\mathbf{p} = \mathbf{f}(0)$

23. $\mathbf{f}(t) = (e^t \operatorname{sen} t, e^t \cos t, 5t)$, en el punto $\mathbf{p} = \mathbf{f}(0)$

24. $\mathbf{f}(t) = (3 \cos t, 1, 2 \operatorname{sen} t)$, en el punto $\mathbf{p} = \mathbf{f}(\pi)$

25. $\mathbf{f}(t) = (-2, 5 \cos^2 t, 10e^{-0.5t})$, en el punto $\mathbf{p} = \mathbf{f}(0)$

26. Determinar los puntos en que la recta tangente a la curva descrita por el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (3t - t^3, 3t^2, 3t + t^3)$ es paralela al plano $3x + y + z = 5$.

27. Hallar el punto en que la recta, imagen del camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t + 1, 3t - 2, 2t - 1)$, está más cerca del origen.

28. Demuestre que el punto $\mathbf{p}$ en que la recta, imagen del camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (at + \alpha, bt + \beta, ct + \gamma)$, está más cerca del origen, es $\mathbf{p} = \mathbf{f}(t_0)$, donde $t_0 = -\frac{a\alpha + b\beta + c\gamma}{a^2 + b^2 + c^2}$.

5.4 Reparametrizaciones

Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino. En varias ocasiones hemos mencionado el hecho de que un camino se puede ver como un objeto matemático caracterizado por dos aspectos: la curva que describe en el espacio $\mathbb{R}^n$ (su traza), que es como la "carretera" por la que circula el punto $\mathbf{f}(t) \in \mathbb{R}^n$, y, por otro lado, la manera como el punto $\mathbf{f}(t)$ recorre tal carretera; más aún, la velocidad a la que el punto realiza el recorrido. Uno de los objetivos fundamentales en el estudio de los caminos $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es el de las propiedades geométricas de las curvas que éstas representan. Estas propiedades no dependen generalmente de la velocidad a la que se recorre la curva, sino solamente de la "curva en sí", es decir, de la "forma geométrica" de la imagen del camino $\mathbf{f}$. Sin embargo, es *vía el camino* $\mathbf{f}$ que se estudian tales propiedades. De hecho, algunos conceptos geométricos importantes en el estudio de la geometría de curvas, se definen por medio de un "recorrido con ciertas características" que se hace en ellas, es decir, por medio de un camino concreto de $\mathbf{f}$ que tenga por imagen la curva en consideración.

En esta sección estudiaremos lo relacionado al parentesco existente entre los diversos caminos que describen una misma curva. Tal parentesco se llama "reparametrización". De manera poco precisa podemos decir que la reparametrización de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ es otro camino $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^n$ con la misma imagen que $\mathbf{f}$. De otro modo, el camino $\bar{\mathbf{f}}$ es una reparametrización del camino $\mathbf{f}$ si la curva descrita por $\bar{\mathbf{f}}$ es la misma que la descrita por $\mathbf{f}$. Esta no será, sin embargo, la definición de reparametrización que aceptaremos. Tendremos que avanzar un poco más en la teoría para afinarle algunos detalles desagradables que admite esta idea general, y establecer así la definición de este concepto que nos servirá.

Antes que nada, un comentario para justificar el término "reparametrización". Ya se había mencionado que a las funciones coordenadas de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ o $\mathbb{R}^3$ se les solía llamar "ecuaciones paramétricas" de la curva que describe el camino. Con tal terminología, a la varibale independiente t se le llama "parámetro". En esta perspectiva entonces, la reparametrización de un camino no es más que un cambio en el parámetro t en las ecuaciones paramétricas de la curva, es decir, un cambio de variable independiente del camino $\mathbf{f}$. Nótese que, viendo el parámetro t como el tiempo en que el punto $\mathbf{f}(t)$ recorre la curva, un cambio en la escala del tiempo representará un cambio de velocidad del desplazamiento del punto $\mathbf{f}(t)$ por la curva. De esta manera, caemos nuevamente en la cuenta de que siendo $\bar{\mathbf{f}}$ un camino que describe la misma curva que $\mathbf{f}$ (una reparametrización de $\mathbf{f}$), la diferencia entre ellos debe estar precisamente en la velocidad de recorrido de la curva.

Tomemos entonces el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ definido en el intervalo I de $\mathbb{R}$. La manera de obtener una reparametrización de $\mathbf{f}$ es simplemente dejar correr el tiempo t en el intervalo I de otra manera. Esto lo podemos lograr definiendo una función $\varphi: J \to I$ de la variable s. Mientras la variable real s recorre el intervalo J, las imágenes $\varphi(s)$ recorrerán el intervalo I; estos serán los valores $t = \varphi(s)$ que tomará $\mathbf{f}$ para asignar los puntos $\mathbf{f}(t) \in \mathbb{R}^n$. La primera condición natural que la función φ debe cumplir es que debe ser sobreyectiva, para asegurar que las imágenes $\varphi(s)$ tomen todos los valores de t en I. Observe entonces que podemos concebir a una reparametrización de $\mathbf{f}$ como una *composición* de $\mathbf{f}$ con φ. Más precisamente, dada la función sobreyectiva $\varphi: J \to I$ definida en el intervalo J de $\mathbb{R}$ y tomando valores en el intervalo I de $\mathbb{R}$, podemos componerla con el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ para formar así la nueva función $\mathbf{f} \circ \varphi: J \subseteq \mathbb{R} \to \mathbb{R}^n$. Siendo $\mathbf{f}$ un camino, esta función es continua. Otra condición natural que surge para la función φ es que sea continua, asegurándose así de que la composición $\mathbf{f} \circ \varphi$ sea una función continua (compuesta de funciones continuas es una función continua) y por lo tanto sea un nuevo camino, el cual, es claro, tiene la misma traza que el camino $\mathbf{f}$ (figura 1)

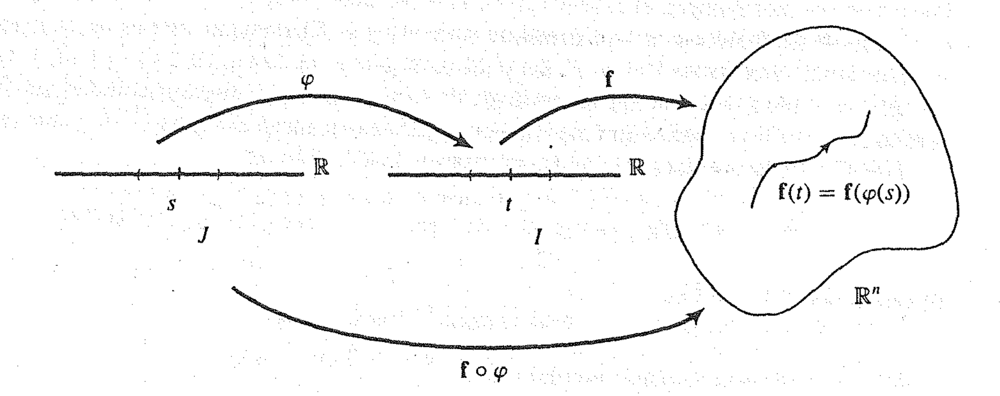


Figura 1. Reparametrización del camino **f**.

Si el camino **f** es diferenciable (respectivamente de clase $\mathscr{C}^1$), nos gustaría que esta propiedad no se perdiera con la reparametrización $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$, de tal manera que también pediremos que la función $t = \varphi(s)$ sea diferenciable (de clase $\mathscr{C}^1$). En tal caso tenemos la composición $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ de funciones diferenciables (de clase $\mathscr{C}^1$), la cual es entonces diferenciable (de clase $\mathscr{C}^1$, respectivamente) y, según la regla de la cadena

$$\bar{\mathbf{f}}'(s) = (\mathbf{f} \circ \varphi)'(s) = \mathbf{f}'(\varphi(s))\varphi'(s),\ s \in J$$

En la fórmula anterior $\varphi'(s)$ es la derivada de la función $\varphi\colon J \to I$ en el punto $s \in J$; por lo tanto es un número real, en tanto que $\mathbf{f}'(\varphi(s))$ es la derivada del camino **f** en el punto $t = \varphi(s)$, el cual es un vector en $\mathbb{R}^n$. Siguiendo la tradición de escribir primero los escalares y luego los vectores en un producto, la fórmula anterior la escribimos como

$$\bar{\mathbf{f}}'(s) = \varphi'(s)\mathbf{f}'(\varphi(s))$$

Esta fórmula descubre la relación entre los vectores velocidad del camino **f** con los correspondientes de su reparametrización $\bar{\mathbf{f}}$: según ella el camino $\bar{\mathbf{f}}$ se mueve en el punto $\bar{\mathbf{f}}(s) \in \mathbb{R}^n$ a una velocidad igual a $\varphi'(s)$ veces la velocidad que llevaría el camino **f** en ese punto (el cual corresponde al punto $t = \varphi(s)$) (figura 2).

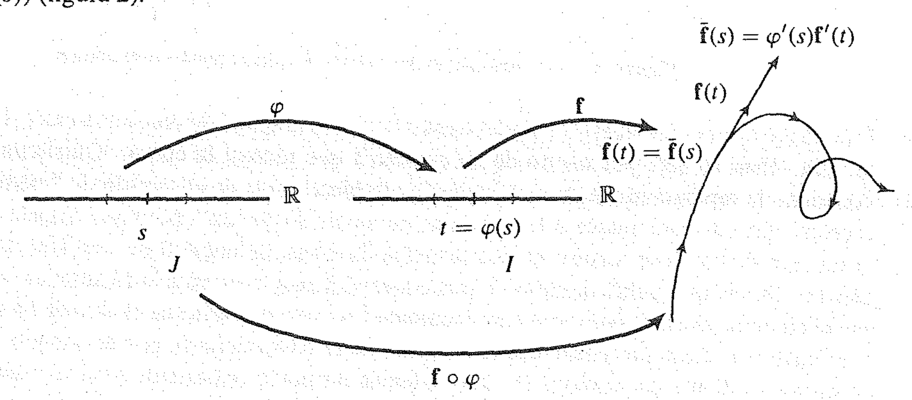


Figura 2. Diagrama de la composición $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$.

Consideremos, por ejemplo, el camino $\mathbf{f}\colon [-1, 1] \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (t, t^2)$. La curva que describe es el arco de parábola $y = x^2$ comprendido entre -1 y 1. Ciertamente este es un camino regular, pues su vector velocidad es $\mathbf{f}'(t) = (1, 2t) \neq (0, 0)\ \forall t \in [-1, 1]$. Sea $\varphi\colon [0, 2\pi] \to [-1, 1]$ la función $t = \varphi(s) = -\cos s$. Esta cumple la condición de sobreyectividad y diferenciabilidad que habíamos pedido para que la composición $\mathbf{f} \circ \varphi\colon [0, 2\pi] \to \mathbb{R}^2$ sea un nuevo camino diferenciable (de hecho, de clase $\mathscr{C}^1$). Esta es entonces la reparametrización $\bar{\mathbf{f}}$ de $\mathbf{f}$, dada por

$$\bar{\mathbf{f}}(s) = (\mathbf{f} \circ \varphi)(s) = \mathbf{f}(\varphi(s)) = \mathbf{f}(-\cos s) = (-\cos s, \cos^2 s),\ s \in [0, 2\pi]$$

El vector velocidad de $\bar{\mathbf{f}}$ es

$$\bar{\mathbf{f}}'(s) = (\operatorname{sen} s, -2 \operatorname{sen} s \cos s)$$

el cual, como dijimos, se puede escribir como

$$\bar{\mathbf{f}}'(s) = (\operatorname{sen} s)(1, -2\cos s) = \varphi'(s)\mathbf{f}'(\varphi(s))$$

Obsérvese que mientras que el camino $\mathbf{f}$ recorre una vez el arco de parábola $y = x^2$ del punto $\mathbf{f}(-1) = (-1, 1)$ al punto $\mathbf{f}(1) = (1, 1)$, tomando para tal recorrido 2 segundos, la reparametrización $\bar{\mathbf{f}}$ efectúa el mismo recorrido de ida y vuelta, comenzando también por el punto $\bar{\mathbf{f}}(0) = (-1, 1)$, yendo de éste al punto $(1, 1) = \bar{\mathbf{f}}(\pi)$ —durante los primeros π segundos— y regresando luego al punto $(-1, 1) = \bar{\mathbf{f}}(2\pi)$ —en los siguientes π segundos— (figura 3).

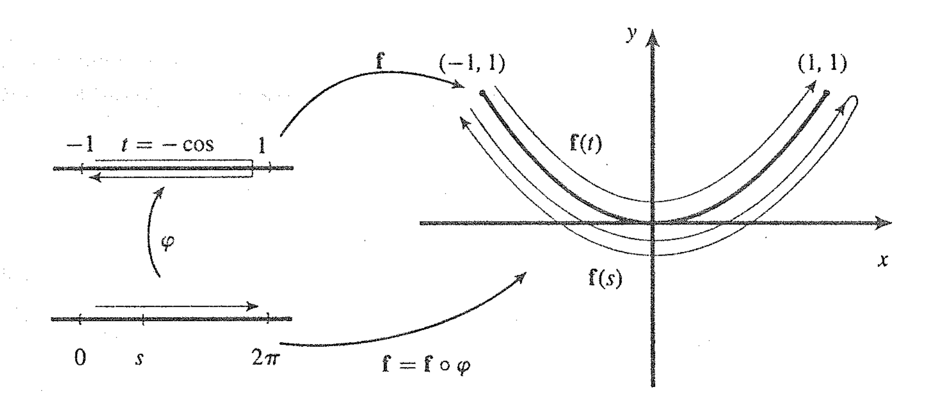


Figura 3. Los recorridos del camino $\mathbf{f}$ y de su reparametrización $\bar{\mathbf{f}}$

En la próxima sección nuestro objetivo será calcular la longitud de una curva entre dos puntos dados de ella. Esto se hará por medio de un camino $\mathbf{f}$ que recorre la curva. Quisiéramos que nuestro concepto de reparametrización que vamos a establecer vaya en el sentido de "maneras distintas de recorrer una curva en cuanto a la velocidad de recorrido de ésta", pero que respete el hecho de que si un arco de curva se recorre de dos maneras distintas, la longitud de éste sólo depende del arco mismo. Desde este punto de vista no nos convendrá aceptar reparametrizaciones de caminos como la del ejemplo anterior, pues mientras el camino $\mathbf{f}$ recorre una sola vez el arco de parábola, el camino $\bar{\mathbf{f}}$ lo hace dos veces (veremos que la longitud de la curva descrita por el camino $\bar{\mathbf{f}}$ es el doble de la correspondiente del camino $\mathbf{f}$). Hay además un punto importante para el cual tendremos que restringir más la función φ con la que obtengamos reparametrizaciones de caminos. Supongamos

que la función φ tiene extremos locales en los puntos s_1 y $s_2 \in J$. Entonces $\varphi'(s_1) = \varphi'(s_2) = 0$. Sean $t_1 = \varphi(s_1)$, $t_2 = \varphi(s_2)$ las correspondientes imágenes en I. Entonces tendríamos

$$\bar{\mathbf{f}}'(s_1) = \varphi'(s_1)\mathbf{f}'(t_1) = 0, \quad \bar{\mathbf{f}}'(s_2) = \varphi'(s_2)\mathbf{f}'(t_2) = 0$$

de modo que si nuestro camino $\mathbf{f}$ original es regular ($\mathbf{f}'(t) \neq 0\ \forall t \in I$), su reparametrización $\bar{\mathbf{f}}$ podría no serlo. La regularidad de un camino es una propiedad importante a conservar en sus reparametrizaciones (figura 4).

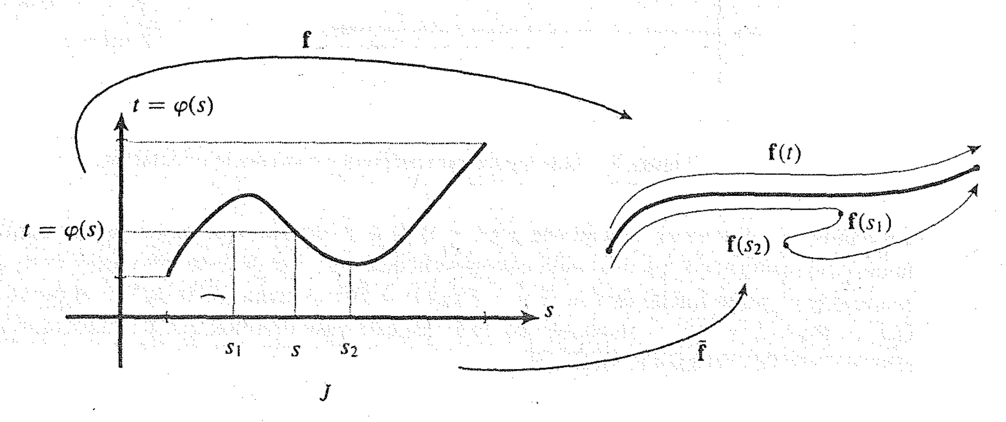


Figura 4. La regularidad del camino $\mathbf{f}$ no la conserva $\bar{\mathbf{f}}$.

Las reparametrizaciones que estaremos interesados en considerar serán aquellas en las que $\varphi'(s) \neq 0$ $\forall s \in J$. Cuando la función φ cumple esta condición adicional es cuando llamamos "reparametrización" a la composición $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$. Con esta nueva condición tenemos garantizado el hecho de que mientras $\mathbf{f}$ recorre la curva de sus imágenes una sola vez, el camino $\bar{\mathbf{f}}$ lo hará también una sola vez (posiblemente con otra velocidad), "sin regresos" como ocurría con $\bar{\mathbf{f}}$ en la figura anterior. Con esto, la regularidad de $\bar{\mathbf{f}}$ (dada la de $\mathbf{f}$) queda automáticamente garantizada. Establezcamos entonces nuestra definición de reparametrización.

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino regular. Sea $\varphi: J \subseteq \mathbb{R} \to I$ una función de clase $\mathscr{C}^1$ sobreyectiva que $\varphi'(s) \neq 0\ \forall s \in J$. Entonces el camino $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi: J \subseteq \mathbb{R} \to \mathbb{R}^n$ se llama *reparametrización* del camino $\mathbf{f}$ (el cual también resulta ser regular). ■

La condición $\varphi'(s) \neq 0\ \forall s \in J$ nos dice que (siendo φ de clase $\mathscr{C}^1$), o $\varphi'(s) > 0\ \forall s \in J$, o $\varphi'(s) < 0\ \forall s \in J$. Escribamos $I = [a, b]$, $J = [c, d]$ y veamos qué sucede en cada uno de los casos.

Si $\varphi'(s) > 0\ \forall s \in J$, entonces φ es una función creciente en J de modo que $\varphi(c) = a$, $\varphi(d) = b$ y así los puntos inicial y final de $\bar{\mathbf{f}}$ coinciden con los respectivos de $\mathbf{f}$, pues $\bar{\mathbf{f}}(c) = (\mathbf{f} \circ \varphi)(c) = \mathbf{f}(\varphi(c)) = \mathbf{f}(a)$ y $\bar{\mathbf{f}}(d) = (\mathbf{f} \circ \varphi)(d) = \mathbf{f}(\varphi(d)) = \mathbf{f}(b)$. Además el vector velocidad $\bar{\mathbf{f}}'(s)$ se obtiene de multiplicar el vector velocidad $\mathbf{f}'(\varphi(s))$ por el escalar positivo $\varphi'(s)$, $s \in J$, de modo que ambos vectores tienen la misma dirección. En este caso, entonces, el camino $\bar{\mathbf{f}}$ recorre la curva descrita por $\mathbf{f}$, en la misma dirección que lo hace $\mathbf{f}$. Diremos que $\bar{\mathbf{f}}$ es una reparametrización de $\mathbf{f}$ que *conserva la orientación.*

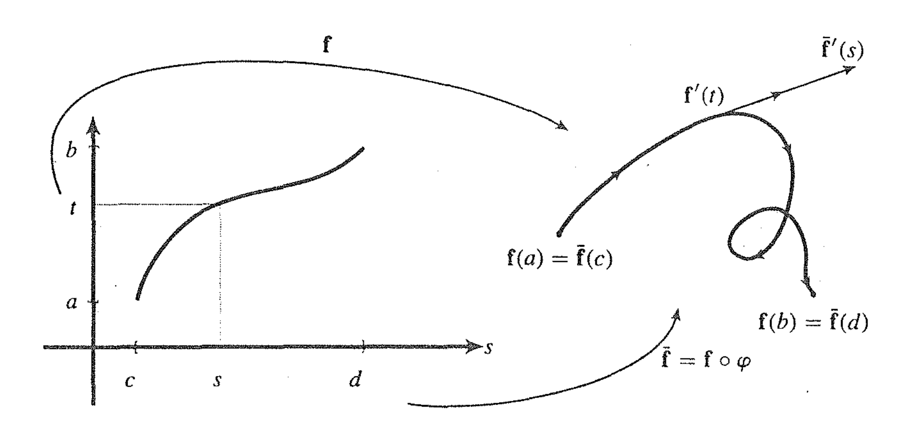


Figura 5. Una reparametrización que conserva la orientación.

Un análisis similar en el caso de que $\varphi'(s) < 0\ \forall s \in J$ nos dice que en tal caso el camino $\bar{\mathbf{f}}$ recorre la curva descrita por $\mathbf{f}$, en dirección contraria la de $\mathbf{f}$. En este caso se tiene $\varphi(c) = b$, $\varphi(d) = a$, de modo que el punto inicial de $\bar{\mathbf{f}}$ es $\bar{\mathbf{f}}(c) = \mathbf{f}(\varphi(c)) = \mathbf{f}(b) =$ punto final de $\mathbf{f}$, y el punto final de $\bar{\mathbf{f}}$ es $\bar{\mathbf{f}}(d) = \mathbf{f}(\varphi(d)) = \mathbf{f}(a) =$ punto inicial de $\mathbf{f}$. En este caso diremos que $\bar{\mathbf{f}}$ es una reparametrización que *invierte la orientación.* (figura 6)

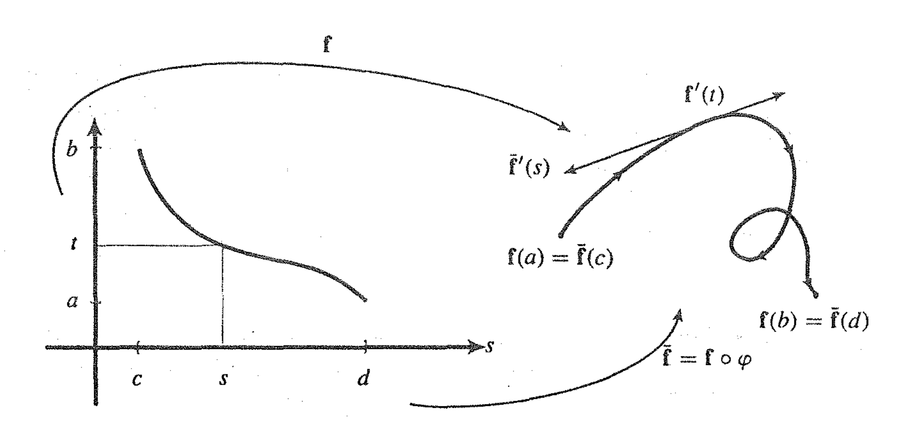


Figura 6. Una reparametrización que invierte la orientación.

Ejemplo 1. Sea $\mathbf{f}\colon [a, b] \to \mathbb{R}^n$ un camino regular. Sea k una constante positiva. Para recorrer la curva descrita por $\mathbf{f}$, k veces más rápido, podemos tomar la función $\varphi\colon \left[0, \frac{b-a}{k}\right] \to [a, b]$ dada por $\varphi(s) = ks + a$. La reparametrización $\bar{\mathbf{f}}\colon \left[0, \frac{b-a}{k}\right] \to \mathbb{R}^n$, $\bar{\mathbf{f}}(s) = (\mathbf{f} \circ \varphi)(s) = \mathbf{f}(ks + a)$ tiene por vector velocidad a

$$\bar{\mathbf{f}}'(s) = \varphi'(s)\mathbf{f}'(\varphi(s)) = k\mathbf{f}'(ks + a), \quad s \in \left(0, \frac{b-a}{k}\right)$$

Si quisiéramos recorrer la curva descrita por **f** con una velocidad en módulo k veces mayor que la correspondiente de **f**, pero en sentido inverso al de **f**, podemos considerar la función $\varphi\colon \left[0, \frac{a-b}{k}\right] \to [a, b]$ dada por $\varphi(s) = ks + b$, donde ahora k es una constante negativa. La reparametrización $\bar{\mathbf{f}}\colon \left[0, \frac{a-b}{k}\right] \to \mathbb{R}^n$, $\bar{\mathbf{f}}(s) = \mathbf{f}(ks + b)$ tiene por vector velocidad a

$$\bar{\mathbf{f}}'(s) = \varphi'(s)\mathbf{f}'(\varphi(s)) = k\mathbf{f}'(ks + b),\ s \in \left[0, \frac{a-b}{k}\right]$$

Un caso particular de esta última situación es cuando $k = -1$. En tal caso el camino $\bar{\mathbf{f}}\colon [0, b-a] \to \mathbb{R}^n$ dado por $\bar{\mathbf{f}}(s) = \mathbf{f}(b - s)$, recorre la curva descrita por **f** con la misma velocidad —en módulo— de **f**, pero en sentido inverso al de **f**. En este caso particular se dice que el camino $\bar{\mathbf{f}}$ es el (camino) *inverso* de **f** y se suele denotar como $-\mathbf{f}$. Una manera alternativa de estudiar este camino, dejándolo definido en el mismo intervalo $[a, b]$ en que está definido **f**, es tomar la función $\varphi\colon [a, b] \to [a, b]$, $\varphi(s) = a + b - s$. Así tendríamos (figura 7)

$$(-\mathbf{f})(s) = \mathbf{f}(\varphi(s)) = \mathbf{f}(a + b - s),\ s \in [a, b]$$

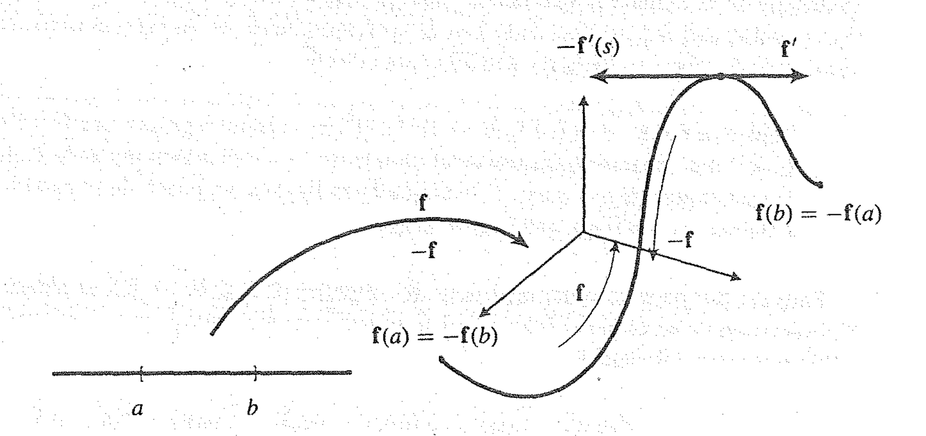


Figura 7. El inverso del camino **f**. ■

Ejemplo 2. Sea $\mathbf{f}\colon [0, 2\pi] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$. La curva que describe **f** es, como sabemos, la circunferencia $x^2 + y^2 = 1$, a partir del punto $(1, 0)$ y en sentido antihorario. Según el ejercicio anterior, tomando $k = 5$, tenemos que el camino $\bar{\mathbf{f}}\colon \left[0, \frac{2\pi}{5}\right] \to \mathbb{R}^2$ dado por $\bar{\mathbf{f}}(s) = \mathbf{f}(5s) = (\cos 5s, \operatorname{sen} 5s)$ recorre la circunferencia $x^2 + y^2 = 1$, 5 veces más rápido que **f**, haciendo dicho recorrido igual que **f**, en sentido antihorario. Poniendo ahora $k = 2$, tenemos que el camino $\bar{\mathbf{f}}\colon [0, \pi] \to \mathbb{R}^2$ dado por $\bar{\mathbf{f}}(s) = \mathbf{f}(-2s + 2\pi) = (\cos(-2s + 2\pi), \operatorname{sen}(-2s + 2\pi)) = (\cos 2s, -\operatorname{sen} 2s)$, recorrerá la curva $x^2 + y^2 = 1$, partiendo de $(1, 0)$ y en sentido horario, con una velocidad del doble (en magnitud) de **f**. El camino inverso de **f** será $-\mathbf{f}\colon [0, 2\pi] \to \mathbb{R}^2$, $(-\mathbf{f})(s) = \mathbf{f}(2\pi - s) = (\cos(2\pi - s), \operatorname{sen}(2\pi - s)) = (\cos s, -\operatorname{sen} s)$. Este camino, como dijimos ya, debe recorrer la curva $x^2 + y^2 = 1$ en 2π segundos (igual que **f**) pero en sentido contrario al de **f** (figura 8).

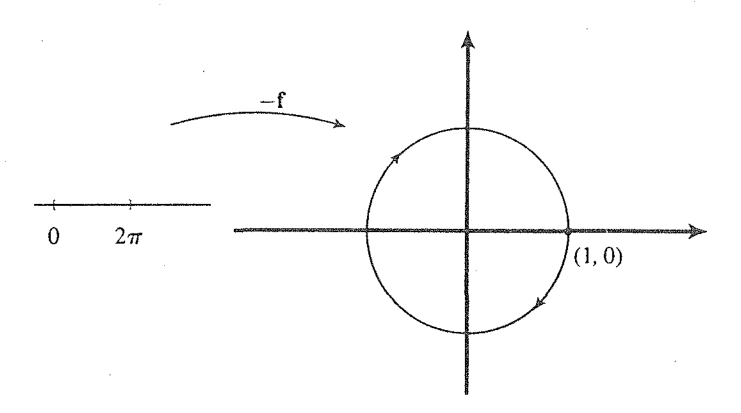


Figura 8. El camino inverso de $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$. ∎

El siguiente resultado, que ya habíamos comentado en la sección anterior, es una consecuencia inmediata de la definición que hemos dado de reparametrización y de la fórmula que relaciona al vector velocidad del camino dado con el correspondiente de su reparametrización. Dejamos su demostración como un ejercicio sencillo para el lector.

Teorema 5.4.1 Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ (o $\mathbb{R}^3$) un camino regular y sea $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi: J \subseteq \mathbb{R} \to \mathbb{R}^2$ (o $\mathbb{R}^3$) una reparametrización de él (donde $\varphi: J \to I$ es la función de la definición). Entonces la recta tangente a la curva C (traza de $\mathbf{f}$) en $\mathbf{f}(t_0)$ (t_0 un punto de I) es la misma que la recta tangente a C en $\bar{\mathbf{f}}(s_0)$, donde $t_0 = \varphi(s_0)$. ∎

Para el caso de un camino regular en $\mathbb{R}^3$, digamos $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$, el plano normal a la curva correspondiente en el punto $\mathbf{f}(t_0)$ está determinado por tal punto y por el vector $\mathbf{f}'(t_0) \in \mathbb{R}^3$. Como habíamos visto, tal plano es

$$x'(t_0)(x - x(t_0)) + y'(t_0)(y - y(t_0)) + z'(t_0)(z - z(t_0)) = 0$$

donde $\mathbf{f}(t) = (x(t), y(t), z(t))$. Sea $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi: J \subseteq \mathbb{R} \to \mathbb{R}^3$ una reparametrización de $\mathbf{f}$. Sea s_0 de modo que $\varphi(s_0) = t_0$. Entonces $\bar{\mathbf{f}}'(s_0) = \varphi'(s_0)\mathbf{f}'(\varphi(s_0)) = \varphi'(s_0)(x'(t_0), y'(t_0), z'(t_0))$, de modo que el plano normal a la curva en $\bar{\mathbf{f}}(s_0) = \mathbf{f}(t_0)$ es

$$\varphi'(s_0)x'(t_0)(x - x(t_0)) + \varphi'(s_0)y'(t_0)(y - y(t_0)) + \varphi'(s_0)z'(t_0)(z - z(t_0)) = 0$$

Como $\varphi'(s_0) \neq 0$, podemos dividir toda la ecuación por $\varphi'(s_0)$ y llegar a la ecuación del plano obtenida con el camino $\mathbf{f}$. Así pues, si una curva C en $\mathbb{R}^3$ es la imagen de un camino regular $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$, las rectas tangentes y planos normales a C se pueden obtener con cualquier reparametrización $\bar{\mathbf{f}}$ de $\mathbf{f}$.

Ejemplo 3. Retomemos el ejemplo 5 de la sección anterior. En él obtuvimos las ecuaciones de la recta tangente y el plano normal a la hélice $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (\cos t, \operatorname{sen} t, t)$ en el punto $\mathbf{f}\left(\frac{\pi}{4}\right)$. Para limitar nuestra atención sólo en un pedazo de la curva alrededor de este punto, restringimos el dominio de $\mathbf{f}$ al intervalo $\left[0, \frac{\pi}{2}\right]$. Sea $\varphi: \left[1, e^{\pi/2}\right] \to \left[0, \frac{\pi}{2}\right]$ la función $\varphi(s) = \ln s$. Entonces

$\bar{\mathbf{f}} = \mathbf{f} \circ \varphi \colon [1, e^{\pi/2}] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = \mathbf{f}(\varphi(s)) = (\cos(\ln s), \operatorname{sen}(\ln s), \ln s)$ es una reparametrización del arco de hélice descrito por $\mathbf{f}$. Para $s = e^{\pi/4}$ tenemos $\bar{\mathbf{f}}(s) = \mathbf{f}\left(\frac{\pi}{4}\right)$. El vector velocidad de $\bar{\mathbf{f}}$ es

$$\bar{\mathbf{f}}'(s) = \left(-\frac{1}{s}\operatorname{sen}(\ln s), \frac{1}{s}\cos(\ln s), \frac{1}{s}\right)$$

que en $s = e^{\pi/4}$ se ve como

$$\begin{aligned}\bar{\mathbf{f}}'(e^{\pi/4}) &= \left(-e^{-\pi/4}\left(\frac{\sqrt{2}}{2}\right), e^{-\pi/4}\left(\frac{\sqrt{2}}{2}\right), e^{-\pi/4}\right)\\ &= e^{-\pi/4}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1\right)\end{aligned}$$

Así que la recta tangente a la curva en $\bar{\mathbf{f}}(e^{\pi/4})$ tiene por ecuaciones paramétricas a

$$\begin{cases} x = \frac{\sqrt{2}}{2} - e^{-\pi/4}\frac{\sqrt{2}}{2}s \\ y = \frac{\sqrt{2}}{2} + e^{-\pi/4}\frac{\sqrt{2}}{2}s \\ z = \frac{\pi}{4} + e^{-\pi/4}s \end{cases} \qquad s \in \mathbb{R}$$

resultado que coincide con el obtenido en el ejemplo 5 de la sección anterior (Las ecuaciones paramétricas de la recta que pasa por (x_0, y_0, z_0) y tiene a $\mathbf{v} = (a, b, c)$ por vector paralelo se pueden ver en general como

$$\begin{cases} x = x_0 + kat \\ y = y_0 + kbt, \\ z = z_0 + kct \end{cases} \qquad t \in \mathbb{R}$$

donde k es cualquier constante no nula, pues al ser $\mathbf{v}$ paralelo a la recta, el vector $k\mathbf{v} = (ka, kb, kc)$ también lo es $(k \neq 0)$). El plano normal pasa por $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{\pi}{4}\right)$ y tiene por vector normal a $\bar{\mathbf{f}}'(e^{\pi/4})$, es entonces

$$-e^{-\pi/4}\frac{\sqrt{2}}{2}\left(x - \frac{\sqrt{2}}{2}\right) + e^{-\pi/4}\frac{\sqrt{2}}{2}\left(y - \frac{\sqrt{2}}{2}\right) + e^{-\pi/4}\left(z - \frac{\pi}{4}\right) = 0$$

o sea

$$z = \frac{\pi}{4} + \frac{\sqrt{2}}{2}(x - y)$$

como se obtuvo también en el ejemplo 5. ■

Quisiéramos ahora llamar la atención a un detalle importante en la definición de reparametrización de un camino y las conclusiones que de ella hemos obtenido. Hemos visto que si el camino $\bar{\mathbf{f}} \colon J \subseteq \mathbb{R} \to \mathbb{R}^n$ es una reparametrización del camino regular $\mathbf{f} \colon I \subseteq \mathbb{R} \to \mathbb{R}^n$, lo cual significa que existe una función $\varphi \colon J \to I$ de clase $\mathscr{C}^1$ y sobreyectiva según la cual $\varphi'(s) \neq 0\ \forall s \in J$ y $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$, entonces los caminos $\bar{\mathbf{f}}$ y $\mathbf{f}$ describen la misma curva, o sea, $\bar{\mathbf{f}}(J) = \mathbf{f}(I)$ (igualdad de conjuntos de puntos en $\mathbb{R}^n$). La pregunta que nos haremos ahora es: suponga que los caminos regulares $\mathbf{f} \colon I \subseteq \mathbb{R} \to \mathbb{R}^n$ y $\bar{\mathbf{f}} \colon J \subseteq \mathbb{R} \to \mathbb{R}^n$ describen la misma curva en $\mathbb{R}^n$. ¿Es $\bar{\mathbf{f}}$ una reparametrización de $\mathbf{f}$? Es decir, ¿existe una función $\varphi \colon J \to I$ de clase $\mathscr{C}^1$ sobreyectiva tal que $\varphi'(s) \neq 0\ \forall s \in I$ y $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$? La respuesta a esta pregunta es NO (necesariamente) y el siguiente par de ejemplos presentan situaciones concretas que la apoyan.

Ejemplo 4. Sea $\mathbf{f}: (-\pi, \pi) \to \mathbb{R}^2$ el camino dado por $\mathbf{f}(t) = (\operatorname{sen} t, \operatorname{sen} 2t)$ y $\bar{\mathbf{f}}: (0, 2\pi) \to \mathbb{R}^2$ el camino $\bar{\mathbf{f}}(s) = (\operatorname{sen} s, \operatorname{sen} 2s)$. Ciertamente estos son caminos regulares y ambos describen la misma curva en $\mathbb{R}^2$ la cual es un "ocho acostado" (recorrido por cada uno de ellos una sola vez) (figura 9).

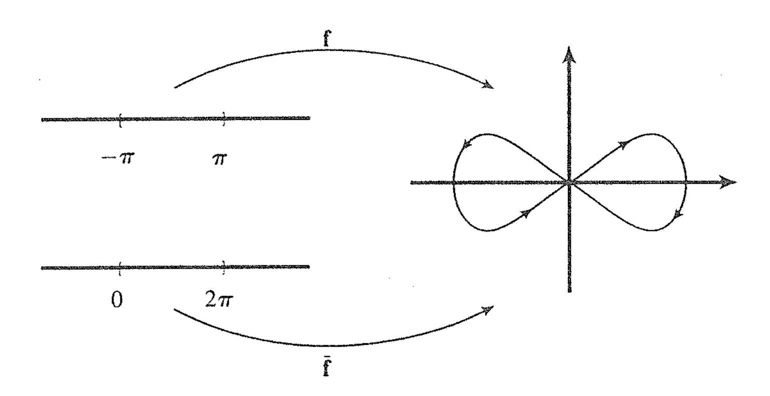


Figura 9. Traza de los caminos $\mathbf{f}$ y $\bar{\mathbf{f}}(t) = (\operatorname{sen} t, \operatorname{sen} 2t)$.

Supongamos, para obtener una contradicción, que $\bar{\mathbf{f}}$ es una reparametrización de $\mathbf{f}$. Entones $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ donde $\varphi: (0, 2\pi) \to (-\pi, \pi)$ es una función de clase $\mathscr{C}^1$ sobreyectiva tal que $\varphi'(s) \neq 0\ \forall s \in (0, 2\pi)$. Obsérvese que $(0, 0) = \bar{\mathbf{f}}(\pi) = \mathbf{f}(\varphi(\pi)) = \mathbf{f}(0)$, de modo que, como $\mathbf{f}$ es inyectiva se tiene $\varphi(\pi) = 0$. Además $\bar{\mathbf{f}}'(s) = (\cos s, 2\cos 2s)$ y $\mathbf{f}'(t) = (\cos t, 2\cos 2t)$, de modo que $\bar{\mathbf{f}}'(\pi) = (-1, 2)$ y $\mathbf{f}'(0) = (1, 2)$. Como $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ se debería tener que $\bar{\mathbf{f}}'(s) = \varphi'(s)\mathbf{f}'(\varphi(s))$, $s \in (0, 2\pi)$. Con $s = \pi$ llegamos a $\bar{\mathbf{f}}'(\pi) = \varphi'(\pi)\mathbf{f}'(\varphi(\pi)) = \varphi'(\pi)\mathbf{f}'(0)$, o sea que $(-1, 2) = \varphi'(\pi)(1, 2)$. Tal número $\varphi'(\pi)$ no existe. Recapacitando esta situación, lo que acontece es que los caminos $\mathbf{f}$ y $\bar{\mathbf{f}}$ pasan por el origen de distintas maneras, como se muestra en la siguiente figura.

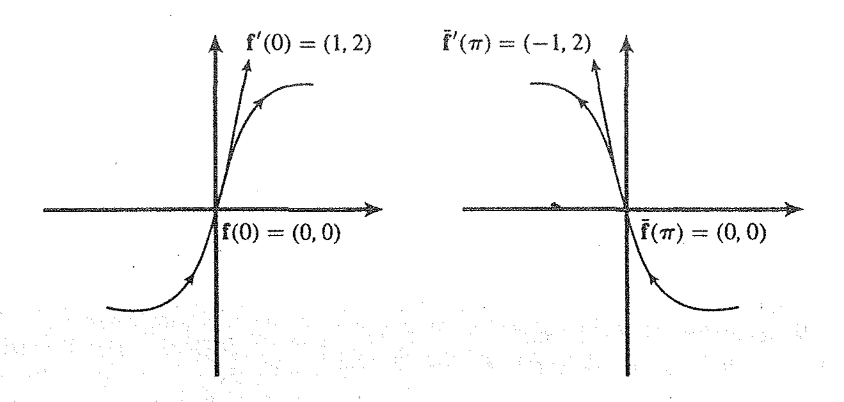


Figura 10. Gráfica de $\mathbf{f}$ y $\mathbf{f}'$ cerca del origen. ■

Ejemplo 5. Consideremos los caminos $\mathbf{f}: [0, 2\pi) \to \mathbb{R}^2$, y $\bar{\mathbf{f}}: [0, 4\pi) \to \mathbb{R}^2$ dados por $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$, $\bar{\mathbf{f}}(s) = (\cos s, \operatorname{sen} s)$. Ambos son caminos regulares que recorren el círculo $x^2 + y^2 = 1$ en sentido antihorario partiendo del punto $(1, 0)$; el camino $\mathbf{f}$ lo recorre una sola vez y el camino $\bar{\mathbf{f}}$ lo hace dos veces. Obsérvese que el camino $\mathbf{f}$ es inyectivo. Supongamos que $\bar{\mathbf{f}}$ es

una reparametrización de $\mathbf{f}$. Entonces $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$. Tenemos $(0, 1) = \bar{\mathbf{f}}(\frac{\pi}{2}) = \mathbf{f}(\varphi(\frac{\pi}{2})) = \mathbf{f}(\frac{\pi}{2})$. Como $\mathbf{f}$ es inyectivo concluimos que $\varphi(\frac{\pi}{2}) = \frac{\pi}{2}$. Análogamente $\bar{\mathbf{f}}(\frac{5\pi}{2}) = \mathbf{f}(\varphi(\frac{5\pi}{2})) = \mathbf{f}(\frac{\pi}{2})$ de donde $\varphi(\frac{5\pi}{2}) = \frac{\pi}{2}$. Entonces $\varphi(\frac{\pi}{2}) = \varphi(\frac{5\pi}{2}) = \frac{\pi}{2}$. Aplicando el Teorema de Rolle a la función φ en el intervalo $[\frac{\pi}{2}, \frac{5\pi}{2}]$, concluimos que debe existir al menos un $s_0 \in (\frac{\pi}{2}, \frac{5\pi}{2})$ tal que $\varphi'(s_0) = 0$, lo cual contradice a la propiedad $\varphi'(s) \neq 0\ \forall s \in [0, 4\pi)$ requerida para que $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ sea una reparametrización de $\mathbf{f}$. ■

Existen resultados que establecen en qué condiciones se puede concluir que el camino $\bar{\mathbf{f}}$ es una reparametrización de $\mathbf{f}$ cuando estos caminos describen la misma curva, pero no ahondaremos en este respecto. Simplemente enfatizamos el hecho estudiado en esta sección de que si $\bar{\mathbf{f}}$ es una reparametrización de $\mathbf{f}$, entonces ambos caminos recorren la misma curva, con distinta velocidad posiblemente, pero que la afirmación recíproca es, en general, falsa.

Ejercicios (Capítulo 5, Sección 4)

En los ejercicios 1–5, considere el camino regular $\mathbf{f}: [-1, 1] \to \mathbb{R}^n$. Diga cuáles de las funciones $\varphi: [-1, 1] \to \mathbb{R}$ dadas , produce una reparametrización $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi: [-1, 1] \to \mathbb{R}^n$ del camino $\mathbf{f}$.

1. $\varphi(s) = s$

2. $\varphi(s) = -s$

3. $\varphi(s) = 2s^2 - 1$

4. $\varphi(s) = 0.5(s - 1)$

5. $\varphi(s) = s^3$

6. Sea $\mathbf{f}: [0, 1] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (t^2 + 1, t^3 + 3t + 2)$. Sea $\varphi: [-1, 1] \to [0, 1]$ una función de clase $\mathscr{C}^1$, sobreyectiva, tal que $\varphi'(s) > 0\ \forall s \in [-1, 1]$. Si $\varphi(0) = 1/2$, $\varphi'(0) = 2$, obtenga el vector velocidad de la reparametrización $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ para $s = 0$.

7. Sea $\mathbf{f}: [-1, 1] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (t^3 - 3t, t^2 + 4t)$. Determine una función $\varphi: [-2, 2] \to [-1, 1]$ sobreyectiva, de clase $\mathscr{C}^1$, tal que $\mathbf{f} \circ \varphi: [-2, 2] \to \mathbb{R}^2$ sea un camino que empiece en el punto $\mathbf{f}(-1) = (2, -3)$, vaya al punto $\mathbf{f}(1) = (-2, 5)$ durante los primeros 2 segundos, y regrese finalmente al punto $\mathbf{f}(-1)$ durante los siguientes 2 segundos. ¿Es $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ una reparametrización de $\mathbf{f}$?

8. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (2\cos t, 3\operatorname{sen} t)$. Dertermine una función $\varphi: [-1, 1] \to [0, 2\pi]$ sobreyectiva, de clase $\mathscr{C}^1$, tal que $\mathbf{g} = \mathbf{f} \circ \varphi: [0, 2\pi] \to \mathbb{R}^2$ recorra 2 veces la elipse $\mathbf{f}([0, 2\pi])$. ¿Es $\mathbf{g}$ una reparametrización de $\mathbf{f}$?

9. Sea $\mathbf{f}: [0, 5] \to \mathbb{R}^3$ el camino $\mathbf{f}(t) = (\operatorname{sen} t, \cos t, t)$. Obtenga una reparametrización $\bar{\mathbf{f}}$ de $\mathbf{f}$, que conserve su orientación y que recorra la traza de $\mathbf{f}$ en la quinta parte de tiempo en que lo hace $\mathbf{f}$.

10. Sea $\mathbf{f}: [0, 1] \to \mathbb{R}^3$ el camino $\mathbf{f}(t) = (t, t, t)$. Obtenga una reparametrización $\bar{\mathbf{f}}$ de $\mathbf{f}$ que invierta su orientación y recorra el segmento de recta $\mathbf{f}([0, 1])$ en 5 segundos.

Para cada uno de los caminos $\mathbf{f}$ dados en los ejercicios 11–15, obtenga una reparametrización $\bar{\mathbf{f}}$ que recorra su imagen de la manera indicada

11. $\mathbf{f}: [0, 2] \to \mathbb{R}^2$, $\mathbf{f}(t) = (3t + 2, t^3 + 3)$. Se quiere que $\bar{\mathbf{f}}$ recorra $\mathbf{f}([0, 2])$ en la misma dirección de $\mathbf{f}$, pero con el doble de velocidad.

12. $\mathbf{f}: [-1, 1] \to \mathbb{R}^2$, $\mathbf{f}(t) = (3\cos t, 3\operatorname{sen} t)$. Se quiere que $\bar{\mathbf{f}}$ recorra $\mathbf{f}([-1, 1])$ en la misma dirección de $\mathbf{f}$, pero a la mitad de velocidad.

13. $\mathbf{f}: [0, 3] \to \mathbb{R}^2$, $\mathbf{f}(t) = (t^2 + t, 4t - 1)$. Se quiere que $\bar{\mathbf{f}}$ recorra $\mathbf{f}([0, 3])$ en dirección contraria a $\mathbf{f}$, y que lo haga en la cuarta parte del tiempo que le toma a $\mathbf{f}$ recorrerlo.

14. $\mathbf{f}: [0, 3] \to \mathbb{R}^3$, $\mathbf{f}(t) = (te^t, e^{-t}, t)$. Se quiere que $\bar{\mathbf{f}}$ recorra $\mathbf{f}([0, 3])$ en dirección contraria a $\mathbf{f}$, y que lo haga en el mismo tiempo de recorrido de $\mathbf{f}$.

15. $\mathbf{f}: [1, 4] \to \mathbb{R}^3$, $\mathbf{f}(t) = (t, 2t, 3t)$. Se quiere que $\bar{\mathbf{f}}$ recorra $\mathbf{f}([1, 4])$ en la misma dirección de $\mathbf{f}$, y que le tome π veces más tiempo en efectuar tal recorrido.

En los ejercicios 16–23, determine una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ que parametrice la curva indicada

16. El eje x, recorrido de izquierda a derecha.

17. El eje y, recorrido de arriba hacia abajo.

18. La recta $y = 2x$, recorrida del tercero al primer cuadrante.

19. La cuarta parte del círculo $x^2 + y^2 = 3$ que se encuentra en el segundo curadrante, recorrido en sentido antihorario.

20. El cuadrado $|x| + |y| = 1$, recorrido en sentido antihorario.

21. El triángulo cuyos vértices son $A = (0, 0)$, $B = (3, 1)$, $C = (2, 5)$, con el recorrido $A \to B \to C \to A$.

22. El segmento de la curva $y = \left|1 - |x|\right|$ comprendido entre $x = -2$ y $x = 2$, recorrido de izquierda a derecha.

23. El segmento de la curva $y = |x^2 - 1|$ comprendido entre $x = -2$ y $x = 2$, recorrido de derecha a izquierda.

En los ejercicios 24–30, determine una función $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ que parametrice la curva indicada.

24. El eje x, recorrido en su dirección positiva.

25. El eje y , recorrido en su dirección negativa.

26. El eje z, recorrido en su dirección negativa.

27. La parte de la recta $y = 2x = 3z$ que se encuentra en el primer octante, comenzando en el origen.

28. La parte de la recta que resulta de la intersección de los dos planos $x+2y-z = 0$, $3x-y+5z = 0$, correspondiente a $z \geq 0$, comenzando en el origen.

29. El círculo que resulta de la intersección del paraboloide $z = x^2 + y^2$ con el plano $z = 4$, comenzando en el punto $(0, 2, 4)$, con el sentido de recorrido $(0, 2, 4) \to (-2, 0, 4) \to (0, -2, 4) \to (2, 0, 4) \to (0, 2, 4)$.

5.5 Longitud de un camino

Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$. En el punto $\mathbf{f}(t_0) \in \mathbb{R}^n$, $t_0 \in I$, tenemos asociado el vector velocidad $\mathbf{f}'(t_0) \in \mathbb{R}^n$ que nos dice en qué dirección se está moviendo el punto $\mathbf{f}(t)$ sobre la curva cuando éste pasa por $\mathbf{f}(t_0)$, y cuya magnitud nos da una estimación numérica de la rapidez con que el punto se mueve. De hecho, en lo subsecuente, llamaremos *rapidez* del camino $\mathbf{f}$ en $\mathbf{f}(t_0)$ al número no negativo $\|\mathbf{f}'(t_0)\|$, la cual distinguiremos del mismo vector velocidad $\mathbf{f}'(t_0)$.

Ejemplo 1. El camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$ que recorre el círculo unitario tiene por vector velocidad a $\mathbf{f}'(t) = (-\operatorname{sen} t, \cos t)$, el cual es distinto en todos los puntos del círculo. Sin embargo, la rapidez a la que se recorre el círculo es siempre constante, pues $\|\mathbf{f}'(t)\| = \sqrt{(-\operatorname{sen} t)^2 + (\cos t)^2} = 1, \forall t \in [0, 2\pi]$. ■

Pensemos el camino $\mathbf{f}: [a, b] \to \mathbb{R}^n$ en términos físicos, como el punto $\mathbf{f}(t)$ moviéndose por la curva descrita por $\mathbf{f}$ con una rapidez $\|\mathbf{f}'(t)\|$; si ésta fuera constante digamos $\|\mathbf{f}'(t)\| = k$ unidades/seg y quisiéramos calcular la longitud total del camino recorrido, todo lo que tendríamos que hacer sería multiplicar a k por el tiempo total empleado en efectuar el recorrido, el cual es $(b - a)$ seg. Así, en este caso la longitud de la curva descrita por $\mathbf{f}$ sería $k(b - a)$ unidades. Esta situación ocurre en el caso del círculo unitario del ejemplo anterior con $k = 1$, $b - a = 2\pi - 0 = 2\pi$, de modo que $k(b - a) = 2\pi$ es la longitud del círculo $x^2 + y^2 = 1$ (recorrido una vez) como ya sabíamos. Sin embargo, en general la rapidez a la que se recorre una curva descrita por un camino es una función del tiempo. Si pensamos "infinitesimalmente", al multiplicar la rapidez $\|\mathbf{f}'(t)\|$ que lleva el punto $\mathbf{f}(t)$ en el instante t, por la diferencial de tiempo dt, tendríamos calculada así la longitud infinitesimal del recorrido en ese instante. Sumando estos productos desde $t = a$ hasta $t = b$ tendríamos calculada la longitud de todo el recorrido hecho por el camino $\mathbf{f}: [a, b] \to \mathbb{R}^n$. Por supuesto que esta suma se hace de manera continua sobre la infinidad de sumandos $\|\mathbf{f}'(t)\|\, dt$, con $t \in [a, b]$. Esta es, como sabemos, la integral

$$\int_a^b \|\mathbf{f}'(t)\|\, dt$$

Este tipo de consideraciones, más bien imprecisas, pero con buena intención, tienen la finalidad de hacer plausible la siguiente definición formal.

Definición. Sea $\mathbf{f}: [a, b] \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$. La longitud de $\mathbf{f}$ entre $t = a$ y $t = b$, denotada por $\ell(\mathbf{f})$, se define como

$$\ell(\mathbf{f}) = \int_a^b \|\mathbf{f}'(t)\| dt$$ ■

Obsérvese que siendo $\mathbf{f}$ un camino de clase $\mathscr{C}^1$ se tiene que $\mathbf{f}'(t)$ es una función continua de t, al igual que $\|\mathbf{f}'(t)\|$ (composición de la función $\varphi(x) = \|x\|$, $x \in \mathbb{R}^n$, con $\mathbf{f}'(t)$, continuas ambas). Así que la integral de la definición anterior existe siempre (funciones continuas son integrables).

Ejemplo 2. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ el camino dado por $\mathbf{f}(t) = (a(t - \operatorname{sen} t), a(1 - \cos t))$, en donde a es un número real positivo dado. La curva que éste describe es el arco de cicloide entre $t = 0$ y $t = 2\pi$ (ver ejemplo 5 de la sección 2). Calculemos la longitud de este arco. Se tiene

$$\|\mathbf{f}'(t)\| = \|(a(1 - \cos t), a \operatorname{sen} t)\| = \sqrt{a^2(1 - \cos t)^2 + a^2 \operatorname{sen}^2 t} = |a|\sqrt{2}\sqrt{1 - \cos t}$$

así que

$$\ell(\mathbf{f}) = \int_a^b \|\mathbf{f}'(t)\|\, dt = \int_0^{2\pi} a\sqrt{2}\sqrt{1-\cos t}\, dt = \sqrt{2}a \int_0^{2\pi} \sqrt{1-\cos t}\, dt$$

$$= 2a \int_0^{2\pi} \sqrt{\operatorname{sen}^2 \frac{t}{2}}\, dt$$

(Obsérvese que $\sqrt{\operatorname{sen}^2 \frac{t}{2}} = |\operatorname{sen} \frac{t}{2}| = \operatorname{sen} \frac{t}{2}$, pues $\operatorname{sen} \frac{t}{2} \geq 0\ \forall t \in [0, 2\pi]$. En general debemos ser muy cuidadosos en respetar los valores absolutos en este tipo de integrales; de otro modo nos podemos llevar sorpresas desagradables como $\int_0^{2\pi} \sqrt{\operatorname{sen}^2 t}\, dt = \int_0^{2\pi} \operatorname{sen} t\, dt = -\cos t\big|_0^{2\pi} = 0$!)

$$= 2a \int_0^{2\pi} \operatorname{sen} \frac{t}{2} dt = -4a \cos \frac{t}{2}\bigg|_0^{2\pi} = 8a$$ ■

Ejemplo 3. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^3$ el camino dado por $\mathbf{f}(t) = (\cos t, \operatorname{sen} t, t)$. La curva que éste describe es la vuelta completa de una hélice, como se muestra en la figura (ver ejemplo 6 de la sección anterior). Calculemos su longitud. Se tiene

$$\|\mathbf{f}'(t)\| = \|(-\operatorname{sen} t, \cos t, 1)\| = \sqrt{\operatorname{sen}^2 t + \cos^2 t + 1} = \sqrt{2}$$

así que

$$\ell(\mathbf{f}) = \int_0^{2\pi} \|\mathbf{f}'(t)\| = \int_0^{2\pi} \sqrt{2}\, dt = 2\sqrt{2}\pi$$

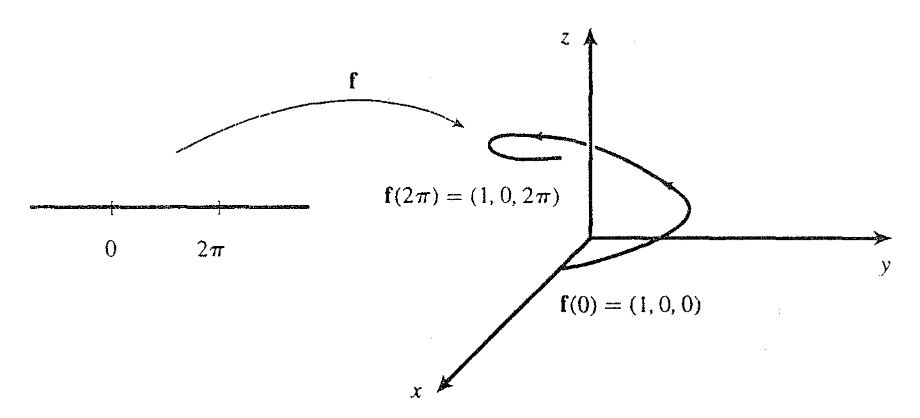


Figura 1. Una vuelta completa de hélice. ■

Ejemplo 4. Dada una curva C en $\mathbb{R}^2$ o $\mathbb{R}^3$ y $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ o $\mathbb{R}^3$ un camino que tiene por traza a C, no es cierto, en general, que la longitud de C sea la longitud del camino $\mathbf{f}$. Por ejemplo si C es una curva cerrada y $\mathbf{f}$ es un camino que recorre a C dos o más veces, la longitud de $\mathbf{f}$ será el número de veces que $\mathbf{f}$ recorre C, multiplicado por la longitud de C. Esta situación ocurre con el camino $\mathbf{f}: [0, 4\pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$ que recorre dos veces el círculo unitario. La longitud de $\mathbf{f}$ será

$$\ell(\mathbf{f}) = \int_0^{4\pi} \|\mathbf{f}'(t)\| = \int_0^{4\pi} 1 dt = 4\pi$$

la cual es dos veces 2π = longitud del círculo unitario $x^2 + y^2 = 1$. ■

Supongamos que $\mathbf{f}: [a, b] \to \mathbb{R}^n$ es un camino regular y que $\bar{\mathbf{f}}: [c, d] \to \mathbb{R}^n$ es una reparametrización de $\mathbf{f}$. Según la discusión presentada en la sección anterior, es de esperarse que la longitud de $\mathbf{f}$ entre $t = a$ y $t = b$ sea igual a la de $\bar{\mathbf{f}}$ entre $s = c$ y $s = d$, pues, "la carretera recorrida por $\mathbf{f}$ y $\bar{\mathbf{f}}$ es la misma", sólo cambia la manera de recorrerla. En efecto, tenemos que $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$, donde $\varphi: [c, d] \to [a, b]$ es de clase $\mathscr{C}^1$, sobreyectiva y $\varphi'(s) \neq 0 \ \forall s \in [c, d]$. Entonces

$$\ell(\bar{\mathbf{f}}) = \int_c^d \|\bar{\mathbf{f}}'(s)\| \, ds = \int_c^d \|\varphi'(s)\mathbf{f}'(\varphi(s))\| \, ds = \int_c^d \|\mathbf{f}'(\varphi(s))\| |\varphi'(s)| \, ds$$

supongamos que $\varphi'(s) > 0 \ \forall s \in [c, d]$. En tal caso tenemos que $\bar{\mathbf{f}}$ es una reparametrización que conserva el sentido de recorrido de $\mathbf{f}$, con $\varphi(c) = a$, $\varphi(d) = b$. Entonces, si $t = \varphi(s)$ en la última integral obtenemos

$$\ell(\bar{\mathbf{f}}) = \int_c^d \|\mathbf{f}'(\varphi(s))\| \varphi'(s) \, ds = \xrightarrow[dt = \varphi'(s) ds]{t = \varphi(s)} = \int_a^b \|\mathbf{f}'(t)\| \, dt = \ell(\mathbf{f})$$

De la misma forma, si $\varphi'(s) < 0, \forall s \in [c, d]$, tenemos que la reparametrización $\bar{\mathbf{f}}$ invierte el sentido, resultando que $\varphi(c) = b$, $\varphi(d) = a$. En tal caso $|\varphi'(s)| = -\varphi'(s)$ y entonces

$$\begin{aligned}
\ell(\bar{\mathbf{f}}) &= \int_c^d \|\mathbf{f}'(\varphi(s))\| (-\varphi'(s)) \, ds = -\int_c^d \|\mathbf{f}'(\varphi(s))\| \varphi'(s) \, ds \\
&= \xrightarrow[dt = \varphi'(s) ds]{t = \varphi(s)} = -\int_b^a \|\mathbf{f}'(t)\| \, dt = \int_a^b \|\mathbf{f}'(t)\| \, dt = \ell(\mathbf{f})
\end{aligned}$$

como asegurábamos que ocurriría.

Ejemplo 5. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (\operatorname{sen} t, \cos t)$. Hemos visto que, para cualquier $n \in \mathbb{N}$, los caminos $\bar{\mathbf{f}}: \left[0, \frac{2\pi}{n}\right] \to \mathbb{R}^2$ dados por $\bar{\mathbf{f}}(s) = (\cos ns, \operatorname{sen} ns)$ son reparametrizaciones de $\mathbf{f}$. Calculemos $\ell(\bar{\mathbf{f}})$

$$\begin{aligned}
\ell(\bar{\mathbf{f}}) &= \int_0^{2\pi/n} \|\bar{f}'(s)\| \, ds = \int_0^{2\pi/n} \|(-n \operatorname{sen} ns, n \cos ns)\| \, ds \\
&= \int_0^{2\pi/n} \sqrt{n^2 \operatorname{sen}^2 ns + n^2 \cos^2 ns} ds = n \int_0^{2\pi/n} ds = n\left(\frac{2\pi}{n}\right) = 2\pi \\
&= \text{perímetro del círculo } x^2 + y^2 = 1 \text{ que describe } \mathbf{f}
\end{aligned}$$

■

Ejemplo 6. Consideremos la curva C en $\mathbb{R}^3$ que se obtiene como intersección de las superficies $y = x^2$ y $z = \frac{2}{3}xy$. Queremos calcular la longitud de ésta desde el origen $(0, 0, 0)$ hasta el punto $\left(1, 1, \frac{2}{3}\right)$. Para usar la fórmula establecida en la definición de la longitud de un camino, debemos disponer de un camino $\mathbf{f}: [a, b] \to \mathbb{R}^3$ de clase $\mathscr{C}^1$ que tenga por imagen a la curva C y que recorra ésta una sola vez. En el ejemplo 8 de la sección 3 nos encontramos con una situación similar. Ahí vimos que resolviendo simultáneamente las ecuaciones que definen a las superficies cuya intersección es la curva C, podemos encontrar parametrizaciones de ella. Por ejemplo, poniendo $x = t$, tenemos que $y = x^2 = t^2$ y $z = \frac{2}{3}xy = \frac{2}{3}t(t^2) = \frac{2}{3}t^3$. Ciertamente el camino $\mathbf{f}(t) = (t, t^2, \frac{2}{3}t^3)$ es un camino de clase $\mathscr{C}^1$ que describe a la curva C. Para $(x =)\ t = 0$ estamos en $(0, 0, 0)$ y para $(x =)\ t = 1$,

estamos en $\left(1, 1, \frac{2}{3}\right)$. Así que la logitud buscada es la del camino $\mathbf{f}: [0, 1] \to \mathbb{R}^3$, $\mathbf{f}(t) = (t, t^2, \frac{2}{3}t^3)$. Tenemos que $\mathbf{f}'(t) = (1, 2t, 2t^2)$ de modo que

$$\|\mathbf{f}'(t)\| = \sqrt{1 + (2t)^2 + (2t^2)^2} = \sqrt{1 + 4t^2 + 4t^4} = \sqrt{(2t^2 + 1)^2} = 2t^2 + 1$$

y entonces

$$\ell(\mathbf{f}) = \int_0^1 \|\mathbf{f}'(t)\|\, dt = \int_0^1 (2t^2 + 1)\, dt = \frac{2}{3} + 1 = \frac{5}{3}$$

En realidad hubiéramos podido poner $x = \varphi(t)$ en la que $\varphi: [a, b] \to [0, 1]$ es cualquier función sobreyectiva de clase $\mathscr{C}^1$ tal que $\varphi'(t) \neq 0\ \forall t \in [a, b]$. En tal caso tendríamos $y = x^2 = (\varphi(t))^2$ y $z = \frac{2}{3}xy = \frac{2}{3}(\varphi(t))^3$. Así, el camino $\mathbf{f}: [a, b] \to \mathbb{R}^3$ dado por $\mathbf{f}(t) = (\varphi(t), (\varphi(t))^2, \frac{2}{3}(\varphi(t))^3)$ es un camino de clase $\mathscr{C}^1$ que describe igualmente a la curva C (en realidad se trata de una reparametrización del camino considerado inicialmente). ■

Ejemplo 7. Considere el camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (\cos^3 t, \operatorname{sen}^3 t)$. La imágen de él es la curva llamada astroide.

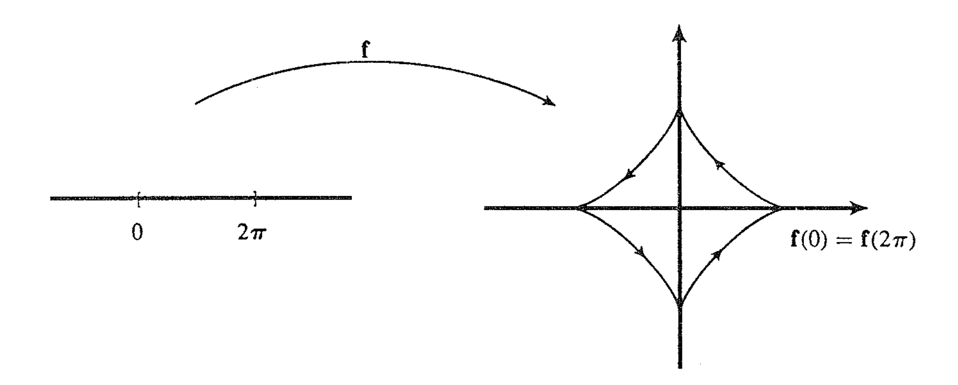


Figura 2. Astroide.

Este no es un camino regular, pero sí es de clase $\mathscr{C}^1$. Calculemos su longitud

$$\begin{aligned}\ell(\mathbf{f}) &= \int_0^{2\pi} \|\mathbf{f}'(t)\|\, dt = \int_0^{2\pi} \|(-3\cos^2 t \operatorname{sen} t, 3\operatorname{sen}^2 t \cos t)\|\, dt \\ &= \int_0^{2\pi} \sqrt{9\cos^4 t \operatorname{sen}^2 t + 9 \operatorname{sen}^4 t \cos^2 t}\, dt = 3\int_0^{2\pi} |\cos t \operatorname{sen} t|\, dt \\ &= \frac{3}{2}\int_0^{2\pi} |\operatorname{sen} 2t|\, dt = \frac{3}{2}(8)\int_0^{\pi/4} \operatorname{sen} 2t\, dt = 6(-\cos 2t)\Big|_0^{\pi/4} = 6\end{aligned}$$

■

Ejemplo 8. Considere una función $\varphi: I \subseteq \mathbb{R} \to \mathbb{R}$ de clase $\mathscr{C}^1$. Podemos calcular la longitud de la gráfica de φ comprendida entre $x = a$ y $x = b$, $a, b \in I$, por medio del camino de clase $\mathscr{C}^1$ $\mathbf{f}: [a, b] \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, \varphi(t))$. Se tiene que

$$\ell(\varphi) = \int_a^b \|\mathbf{f}'(t)\|\, dt = \int_a^b \|(1, \varphi'(t))\|\, dt = \int_a^b \sqrt{1 + (\varphi'(t))^2}\, dt$$

Por ejemplo, si $\varphi: \mathbb{R} \to \mathbb{R}$ es la función $\varphi(x) = \cosh x$, tenemos que la longitud de la gráfica (llamada catenaria) entre $x = a$ y $x = b$ es

$$\begin{aligned}\ell(\varphi) &= \int_a^b \sqrt{1 + (\varphi'(t))^2} dt = \int_a^b \sqrt{1 + \operatorname{senh}^2 t} dt = \int_a^b \sqrt{\cosh^2 t} dt \\ &= \int_a^b \cosh t dt = \operatorname{senh} t\big|_a^b = \operatorname{senh} b - \operatorname{senh} a\end{aligned}$$

(Recuerde que $\operatorname{senh} t = \frac{e^t - e^{-t}}{2}$ y $\cosh t = \frac{e^t + e^{-t}}{2}$ y se cumple que $\cosh^2 t - \operatorname{senh}^2 t = 1$, $\forall t \in \mathbb{R}$)

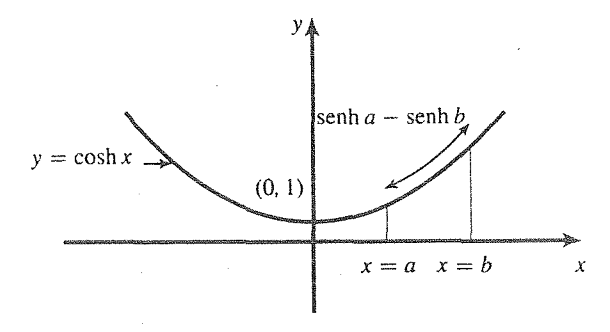


Figura 3. Longitud de la catenaria $y = \cosh x$ entre $x = a$ y $x = b$. ■

Sean $\mathbf{f}: [a, b] \to \mathbb{R}^n$ y $\mathbf{g}: [c, d] \to \mathbb{R}^n$ dos caminos tales que $\mathbf{f}(b) = \mathbf{g}(c)$, es decir, el punto donde termina el camino $\mathbf{f}$, es el punto donde comienza el camino $\mathbf{g}$. Podríamos pensar en un camino "$\mathbf{f}$ seguido de $\mathbf{g}$", que recorre el camino $\mathbf{f}$ y continúa recorriendo luego el camino $\mathbf{g}$. Para dejar este nuevo camino definido en un intervalo podemos reparametrizar el camino $\mathbf{g}$ de modo que éste

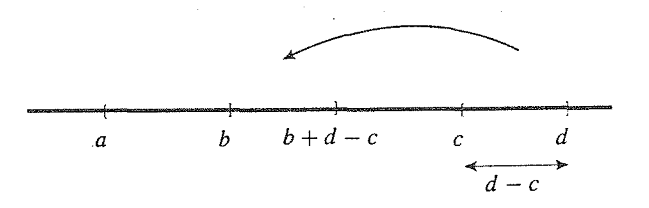
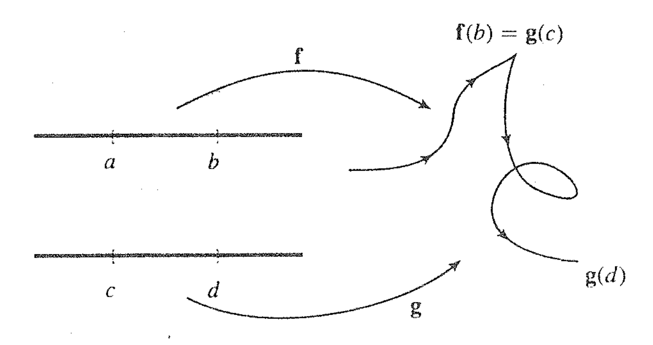


Figura 4. Arreglo de los intervalos $[a, b]$, $[c, d]$.

quede definido en el intervalo $[b, b+d-c]$ por medio de la función $\varphi: [a, b+d-c] \to [c, d]$, $\varphi(s) = s - b + c$ (figura 4).

Denotaremos este nuevo camino como $\mathbf{f} + \mathbf{g}$ (advertencia: hemos usado anteriormente esta notación para la función $(\mathbf{f}+\mathbf{g})(t) = \mathbf{f}(t) + \mathbf{g}(t)$. Esto, sin embargo, no debe provocar confusión, atendiendo al contexto del problema). Este camino es entonces $\mathbf{f} + \mathbf{g}: [a, b+d-c] \to \mathbb{R}^n$

$$(\mathbf{f}+\mathbf{g})(t) = \begin{cases} \mathbf{f}(t) & \text{si } t \in [a, b] \\ \mathbf{g}(t-b+c) & \text{si } t \in [b, b+d-c] \end{cases}$$



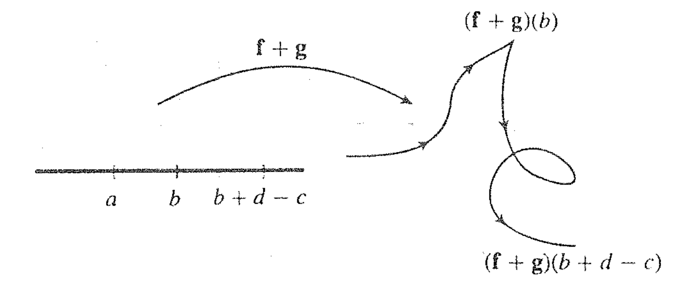


Figura 5. El camino $\mathbf{f} + \mathbf{g}$.

Si tanto $\mathbf{f}$ como $\mathbf{g}$ son caminos de clase $\mathscr{C}^1$, es claro que el camino $\mathbf{f} + \mathbf{g}$ puede no serlo. Sin embargo, podemos definir la longitud de este último camino entre $t = a$ y $t = b + d - c$ como

$$\ell(\mathbf{f}+\mathbf{g}) = \ell(\mathbf{f}) + \ell(\mathbf{g}) = \int_a^b \|\mathbf{f}'(t)\|\,dt + \int_c^d \|\mathbf{g}'(t)\|\,dt$$

Ejemplo 9. Considere los caminos $\mathbf{f}: [0, \pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = (-\cos t, \operatorname{sen} t)$ y $\mathbf{g}: [1, 3] \to \mathbb{R}^2$, $\mathbf{g}(t) = (t, t-1)$. Se tiene $\mathbf{f}(\pi) = (1, 0) = \mathbf{g}(1)$. Formemos entonces el camio $\mathbf{f} + \mathbf{g}$. Este es $\mathbf{f} + \mathbf{g}: [0, \pi + 2] \to \mathbb{R}^2$

$$(\mathbf{f}+\mathbf{g})(t) = \begin{cases} (-\cos t, \operatorname{sen} t) & \text{si } t \in [0, \pi] \\ (t - \pi + 1, t - \pi) & \text{si } t \in [\pi, \pi + 2] \end{cases}$$

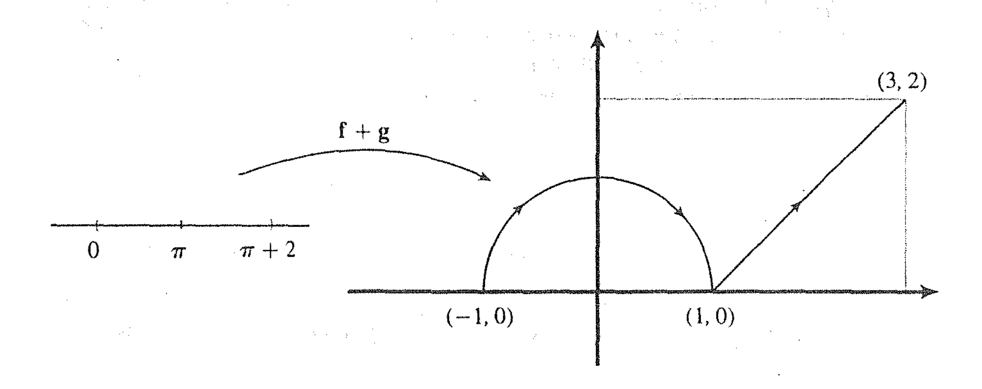


Figura 6. El camino **f** + **g** del ejemplo 9.

Nótese que **f** y **g** son caminos de clase $\mathscr{C}^1$, pero **f** + **g** no lo es, pues

$$\lim_{t\to\pi^+}(\mathbf{f}+\mathbf{g})'(t) = \lim_{t\to\pi^+}\mathbf{f}'(t) = \lim_{t\to\pi^+}(\operatorname{sen} t, \cos t) = (0, -1)$$
$$\lim_{t\to\pi^-}(\mathbf{f}+\mathbf{g})'(t) = \lim_{t\to\pi^-}\mathbf{g}'(t) = \lim_{t\to\pi^-}(1, 1) = (1, 1)$$

así que $\lim_{t\to\pi^-}(\mathbf{f}+\mathbf{g})'(t)$ no existe, y entonces la derivada $(\mathbf{f}+\mathbf{g})'(t)$ es discontinua en $t = \pi$. Sin embargo, podemos calcular la longitud del camino **f** + **g** entre $t = 0$ y $t = \pi + 2$ como

$$\begin{aligned}\ell(\mathbf{f}+\mathbf{g}) &= \ell(\mathbf{f}) + \ell(\mathbf{g}) = \int_0^{\pi} \|\mathbf{f}'(t)\|\,dt + \int_1^3 \|\mathbf{g}'(t)\|\,dt \\ &= \int_0^{\pi} \|(\operatorname{sen} t, \cos t)\|\,dt + \int_1^3 \|(1, 1)\|\,dt = \int_0^{\pi} dt + \int_1^3 \sqrt{2}\,dt \\ &= \pi + 2\sqrt{2}\end{aligned}$$

■

Tenemos entonces que el camino **f** + **g** obtenido al unir los caminos **f** y **g** de clase $\mathscr{C}^1$, puede no ser de clase $\mathscr{C}^1$, pero es posible partir el intervalo donde está definido de modo que restringiendo a cada uno de los subintervalos de la partición, quede de clase $\mathscr{C}^1$. De hecho así es como se construye el camino **f** + **g**. En términos más generales, diremos que un camino $\mathbf{f}: [a, b] \to \mathbb{R}^n$ es *seccionalmente de clase* $\mathscr{C}^1$ si se da una partición del intervalo $[a, b]$ digamos

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b$$

de modo que restringiendo el camino a cada subintervalo $[t_{i-1}, t_i]$ $i = 1, 2, \ldots, n$, se obtenga un camino de clase $\mathscr{C}^1$. En tal caso podemos definir la longitud del camino **f** entre $t = a$ y $t = b$ como

$$\ell(\mathbf{f}) = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \|\mathbf{f}'(t)\|\,dt$$

donde la derivada $\mathbf{f}'$ se calcula en la función **f** restringiendo al subintervalo $[t_{i-1}, t_i]$. Es decir, la longitud del camino **f** es la suma de longitudes de ese camino en los subintervalos en que **f** "se porta bien" (de clase $\mathscr{C}^1$).

Ejemplo 10. El camino $\mathbf{f}: [-\alpha, \alpha] \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (t, |t|)$ no es de clase $\mathscr{C}^1$. Sin embargo podemos partir el intervalo $[-\alpha, \alpha]$ como $-\alpha = t_0 < 0 < t_2 = \alpha$, de modo que $\mathbf{f}$ restringido a $[-\alpha, 0]$ es $\mathbf{f}_1(t) = (t, -t)$ que es de clase $\mathscr{C}^1$, y restringido a $[0, \alpha]$ es $\mathbf{f}_2(t) = (t, t)$, que también es de clase $\mathscr{C}^1$. Así entonces f es un camino seccionalmente $\mathscr{C}^1$. Su longitud es pues

$$\begin{aligned}\ell(\mathbf{f}) &= \int_{-\alpha}^{0} \|\mathbf{f}_1'(t)\|\, dt + \int_{0}^{\alpha} \|\mathbf{f}_2'(t)\|\, dt \\ &= \int_{-\alpha}^{0} \|(1, -1)\|\, dt + \int_{0}^{\alpha} \|(1, 1)\|\, dt \\ &= \int_{-\alpha}^{0} \sqrt{2}\, dt + \int_{0}^{\alpha} \sqrt{2}\, dt = \sqrt{2}\alpha + \sqrt{2}\alpha = 2\sqrt{2}\alpha\end{aligned}$$ ■

Ejercicios (Capítulo 5, Sección 5)

En los ejercicios 1–5, calcule la longitud de la gráfica de las funciones $y = \varphi(x)$, comprendida en los intervalos indicados (ver ejemplo 8).

1. $y = 5x + 1, 0 \leq x \leq 3$

2. $y = \ln x, 1 \leq x \leq 5$

3. $y = x^{3/2}, 0 \leq x \leq 1$

4. $y = 3\cosh \frac{x}{3}, 0 \leq x \leq 2$

5. $y = \ln \cos x, 0 \leq x \leq \pi/3$

En los ejercicios 6–10, calcule la longitud de los caminos dados.

6. $\mathbf{f}: [0, 3] \to \mathbb{R}^2, \mathbf{f}(t) = (t, 5t + 1)$ (ver ejercicio 1)

7. $\mathbf{f}: [0, \sqrt{2}] \to \mathbb{R}^2, \mathbf{f}(t) = (8t^3, 6t^2 - 3t^4)$

8. $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2, \mathbf{f}(t) = (\cos^4 t, \operatorname{sen}^4 t)$

9. $\mathbf{f}: [0, 1] \to \mathbb{R}^2, \mathbf{f}(t) = (\cosh^2 t, \operatorname{senh}^2 t)$

10. $\mathbf{f}: [0, 1] \to \mathbb{R}^3, \mathbf{f}(t) = (\cosh t, \operatorname{senh} t, t)$

11. Considere el camino $\lambda: [0, 2\pi] \to \mathbb{R}^2$, $\lambda(t) = (a \cos t, b \operatorname{sen} t)$, cuya imágen es (como sabemos), la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Demuestre que la longitud L de esta curva viene dada por la integral

$$L = 4b \int_{0}^{\pi/2} \sqrt{1 - e^2 \operatorname{sen}^2 t}\, dt$$

donde $e = \frac{1}{b}\sqrt{b^2 - a^2}$ (este número es la excentricidad de la elipse). Para obtener el valor de la integral anterior, haciendo uso del teorema fundamental del cálculo, habría que determinar primeramente la integral indefinida correspondiente. Sin embargo, ésta última (puede

demostrarse que) no tiene solución en términos de funciones elementales. De hecho, la integral que aparece en la expresión de L es un caso particular de integrales de la forma

$$\int_0^{\pi/2} \sqrt{1 - k^2 \operatorname{sen}^2 t}\, dt$$

donde $0 < k < 1$, llamadas *integrales elípticas de segunda clase*. Estas integrales juegan un papel muy importante en algunas partes de la matemática, y se tienen referencias de valores de ellas para diferentes valores de k.

12. Considere el camino $\mathbf{f}: [t_0, +\infty) \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (e^{-t} \cos t, e^{-t} \operatorname{sen} t)$.

a. Demuestre que la imagen de $\mathbf{f}$ cruza una infinidad de veces cada uno de los ejes coordenados. Más aún, (demuestre que) si $t, t' \in [t_0, +\infty)$ son tales que $t > t'$ y $\mathbf{f}(t)$ y $\mathbf{f}(t')$ se encuentran en el eje x (o en el eje y), entonces $\|\mathbf{f}(t)\| < \|\mathbf{f}(t')\|$. Interprete vía geométrica este hecho.

b. Demuestre que cuando t tiende a infinito, la imagen de $\mathbf{f}$ tiende al origen de coordenadas (¿de qué manera?).

c. Considere la longitud de $\mathbf{f}$ entre $t = t_0$ y $t = t_1$. Denótela como $L_{[t_0, t_1]}$. Defina la *longitud total* de $\mathbf{f}$ como el límite $\lim_{t_1 \to \infty} L_{[t_0, t_1]}$. Demuestre que la longitud total de $\mathbf{f}$ es finita. Calcúlela. A la traza de $\mathbf{f}$ se le llama *espiral logarítmica*.

13. Considere los caminos $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}: [\alpha, \beta] \to \mathbb{R}^3$ dados por $\mathbf{f}(t) = (t, \varphi(t), a)$, $\mathbf{g}(t) = (t, b, \varphi(t))$, $h(t) = (c, t, \varphi(t))$, donde a, b, c son números reales dados y $\varphi: [\alpha, \beta] \to \mathbb{R}$ es una función de clase $\mathscr{C}^1$. Demuestre que la longitud de los caminos $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$ es igual a la longitud de la gráfica de la función φ. Interprete geométricamente.

14. ¿Cuál es la longitud de un camino constante $\mathbf{f}: \mathbb{R} \to \mathbb{R}^n$, $\mathbf{f}(t) = \mathbf{v}$?

15. Determine la longitud del camino $\mathbf{f}: [\alpha, \beta] \to \mathbb{R}^n$ dado por $\mathbf{f}(t) = (a_1 t + b_1, a_2 t + b_2, \ldots, a_n t + b_n)$. Demuestre que ésta se puede escribir como $\|\mathbf{f}(\alpha) - \mathbf{f}(\beta)\|$. Interprete geométricamente.

16. Sean $\mathbf{p}$ y $\mathbf{q}$ dos puntos en $\mathbb{R}^3$. El objetivo de este ejercicio es probar que la distancia más corta entre estos dos puntos es $\|\mathbf{q} - \mathbf{p}\|$ = longitud del camino $\mathbf{f}: [0, 1] \to \mathbb{R}^3$, $\mathbf{f}(t) = t\mathbf{q} + (1 - t)\mathbf{p}$ (es decir, la distancia más corta entre $\mathbf{p}$ y $\mathbf{q}$ es la longitud de la recta que une a $\mathbf{p}$ con $\mathbf{q}$). Sea $\mathbf{f}: [a, b] \to \mathbb{R}^3$ un camino de clase $\mathscr{C}^1$, tal que $\mathbf{f}(a) = \mathbf{p}$, $\mathbf{f}(b) = \mathbf{q}$.

a. Sea $\mathbf{v} \in \mathbb{R}^3$ un vector unitario. Considere la función $\varphi: [a, b] \to \mathbb{R}$ dada por $\varphi(t) = \mathbf{f}'(t) \cdot \mathbf{v}$. Demuestre que

$$\int_a^b \varphi(t)\, dt = (\mathbf{q} - \mathbf{p}) \cdot \mathbf{v}$$

b. Use la desigualdad de Cauchy-Schawrz con los vectores $\mathbf{f}'(t)$ y $\mathbf{v}$ para demostrar que

$$\int_a^b \varphi(t)\, dt \leq \int_a^b \|\mathbf{f}'(t)\|\, dt$$

c. Considere el vector unitario $\mathbf{v} = \frac{\mathbf{q} - \mathbf{p}}{\|\mathbf{q} - \mathbf{p}\|}$ para demostrar, usando los

$$\|\mathbf{q} - \mathbf{p}\| \leq \text{longitud del camino } \mathbf{f}$$

d. Concluya que la distancia más corta entre $\mathbf{p}$ y $\mathbf{q}$ es la de la línea recta que une estos dos puntos.

17. (En un tono menos serio: un problema intergaláctico). Era el 19 de diciembre del año 2015. El capitán Marcello se disponía a realizar su misión número $\pi^2 \times 10^2$. A bordo de su nave "Apolo-Tepepan", escuchaba la cuenta regresiva para dar inicio a su vuelo por el espacio. En esta ocasión el capitán Marcello debía dirigirse a un punto en el espacio $\mathbb{R}^3$ con coordenadas $(50, 70, 100)$ —las unidades están medidas en miles de kilómetros—, teniendo el punto de arranque en el origen de coordenadas. Después de que su equipo de trabajo había resuelto el ejercicio anterior, y convencidos plenamente de que la distancia más corta entre el punto de despegue y el punto objetivo es la línea recta, planearon un viaje descrito matemáticamente por el camino $\mathbf{f}: [0, 1] \to \mathbb{R}^3$, $\mathbf{f}(t) = (50t, 70t, 100t)$, en donde t está medido en cientos de horas de vuelo. La computadora ha indicado el paso por tres "zonas de alta turbulencia", en donde se deben extremar las precauciones de vuelo, como por ejemplo, dejar de tomar café y no ir al sanitario. Estas zonas estaban registradas en forma matemática, con ecuaciones que describían las regiones que éstas ocupaban en el espacio. Así, bajo el título de "PELIGRO, ZONAS DE ALTA TURBULENCIA", apareció en la pantalla de la computadora de la nave la siguiente información:

$$\text{ZONA UNO:}\{(x, y, z)|x^2 + y^2 + z^2 - 16x - 24y - 30z + 424 \leq 0\}$$
$$\text{ZONA DOS:}\{(x, y, z)|x^2 + y^2 + z^2 - 36x - 50y - 80z + 2533 \leq 0\}$$
$$\text{ZONA TRES:}\{(x, y, z)|x^2 + y^2 + z^2 - 60x - 80y - 110z + 5489 \leq 0\}$$

"Es fácil ver que se trata de zonas esféricas", dijo para sí mismo el capitán Marcello, segundos después de ver las ecuaciones en pantalla. Tomó papel y lápiz, y se dispuso a hacer los cálculos de los tiempos en que entraría y saldría de cada una de las zonas de turbulencia, así como de la cantidad de kilómetros que recorrería en cada una de ellas. Con un gesto de tranquilidad, pudo ver que no pasaría más de 20 horas, de las 100 horas de vuelo, dentro de zonas de turbulencia.

Reproduzca los cálculos que hizo el capitán Marcello, indicando tiempo de entrada y tiempo de salida a cada una de las tres zonas, así como la distancia recorrida en cada una de ellas.

Después de 25 horas de vuelo, habiendo dejado atrás la primera zona de turbulencia, apareció en la pantalla una información de última hora: justo en el momento del despegue, se detectó una explosión en el espacio, registrada en el punto de coordenadas $(35, 50, 80)$. Este tipo de fenómenos espaciales eran de gran peligro, pues si se entraba en contacto con la región de las ondas expansivas de la explosión en menos de 150 horas de que ocurrió ésta, las probabilidades de que la nave sufriera transtornos serios en las turbinas, eran de alrededor de $0.5e^{1/2}$. Se sabe que las ondas expansivas eran esferas concéntricas (con centro en el punto de la explosión) cuyo radio crecía de manera proporcional al tiempo transcurrido desde la explosión; es decir, el radio de las esferas de las ondas expansivas era del tipo $r = kt$, en donde k es una constante positiva. ¿Hasta qué valor de k la nave del capitán Marcello no cruzaría la región de ondas expansivas?

18. Para adornar la entrada al Palacio de Mathingham (cuya historia se relatará en el ejemplo 14 de la sección 7 del capítulo 7, ¡no se la pierda!), se tiene pensado colocar tres cadenas colgantes. Las cadenas serán de oro macizo y colgarán cada una de ellas de dos postes consecutivos de los cuatro postes que se colocarán al frente del jardín principal. La altura de cada uno de estos postes es de 1 mt y la distancia entre dos postes consecutivos es de 2 mt. Para cada una de las dos cadenas laterales se quiere que la mínima distancia de ellas al piso sea de 0.5 mt, mientras que para la cadena central se quiere que esta distancia sea de 0.25 mt. Se sabe que la curva que forma una cadena colgante es del tipo $y = a \cosh kx$. Poniendo el eje y en el eje de simetría de cada una de las cadenas, demuestre que las dos cadenas laterales pueden ser representadas por la ecuación

$y = 0.5\cosh((\text{arccosh } 2)x)$, y la cadena central por la ecuación $y = 0.25\cosh((\text{arccosh } 4)x)$. Demuestre que la longitud total de cadena de oro que será necesaria para este fastuoso adorno del Palacio de Mathingham es de (aproximadamente) 7.26 mt. (Use integración numérica para obtener el valor de las integrales que le aparecerán en el cálculo de las longitudes de las catenarias involucradas).

5.6 Reparametrización por longitud de arco

En esta sección veremos la posibilidad de reparametrizar un camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$, digamos que $\bar{\mathbf{f}}\colon [a, b] \to \mathbb{R}^n$ es la reparametrización, de modo que la rapidez a que $\bar{\mathbf{f}}$ recorre la curva C = imagen de $\mathbf{f}$ = imagen de $\bar{\mathbf{f}}$, sea constante e igual a la unidad. Es decir, que $\|\bar{\mathbf{f}}'(t)\| = 1$, $\forall t \in [a, b]$. Obsérvese que en tal caso la longitud del camino $\bar{\mathbf{f}}$ entre $t = a$ y $t = b$ es

$$\ell(\bar{\mathbf{f}}) = \int_a^b \|\bar{\mathbf{f}}'(t)\|\, dt = \int_a^b 1\, dt = b - a$$

así que $\bar{\mathbf{f}}$ será una reparametrización tal que la longitud de la curva que describe es igual al tiempo que emplea en recorrerla.

Por ejemplo, el camino $\mathbf{f}\colon [0, 2\pi] \to \mathbb{R}^2$, dado por $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$, que recorre una vez el círculo unitario $x^2 + y^2 = 1$, tiene las características menicionadas: la rapidez con que se recorre el círculo es $\|\mathbf{f}'(t)\| = \|(-\operatorname{sen} t, \cos t)\| = \sqrt{\operatorname{sen}^2 t + \cos^2 t} = 1$ $\forall t \in [0, 2\pi]$ y, consecuentemente, la longitud de la curva descrita por $\mathbf{f}$ es 2π = tiempo que emplea en recorrerla.

Consideremos un camino regular $\mathbf{f}\colon [a, b] \to \mathbb{R}^n$ y supongamos un punto $\mathbf{p} = \mathbf{f}(t_0) \in \mathbb{R}^n$ de la curva que describe (t_0 algún punto en $[a, b]$). Siendo $\mathbf{f}$ de clase $\mathscr{C}^1$ podemos calcular, para cada $t \in [a, b]$, la longitud de $\mathbf{f}$ entre $t = t_0$ y t. Esta es claramente una función de t, designada por $\psi(t)$. Entonces

$$\psi(t) = \int_{t_0}^t \|\mathbf{f}'(u)\|\, du$$

(Aceptamos que si $t < t_0$ (i.e. el punto $\mathbf{f}(t)$ está antes —en el recorrido de $\mathbf{f}$— que $\mathbf{p} = \mathbf{f}(t_0)$), entonces la longitud del camino es negativa). Según el teorema fundamental del cálculo tenemos

$$\psi'(t) = \|\mathbf{f}'(t)\|, \quad t \in [a, b]$$

Como nuestro camino $\mathbf{f}$ es regular, entonces $\psi'(t) = \|\mathbf{f}'(t)\| \neq 0$ $\forall t \in [a, b]$. De hecho se tiene $\psi'(t) = \|\mathbf{f}'(t)\| > 0$ $\forall t \in [a, b]$. Además, es claro que ψ es una función de clase $\mathscr{C}^1$, pues su derivada $\psi'(t) = \|\mathbf{f}'(t)\|$ es continua.

Sea $[c, d] \subset \mathbb{R}$ el intervalo en el que la función ψ manda las imágenes $\psi(t)$, $t \in [a, b]$. Por ejemplo, si hubiéramos tomado $\mathbf{p} = \mathbf{f}(a)$, entonces la función $\psi\colon [a, b] \to \mathbb{R}$

$$\psi(t) = \int_a^t \|\mathbf{f}'(u)\|\, du$$

manda el intervalo $[a, b]$ al intervalo $[0, \ell(\mathbf{f})]$ sobreyectivamente, donde $\ell(\mathbf{f})$ es la longitud de $\mathbf{f}$ entre $t = a$ y $t = b$.

Sabemos, de nuestro primer curso de cálculo, que siendo $\psi: [a, b] \to [c, d]$ una función de clase $\mathscr{C}^1$ tal que $\psi'(t) \neq 0\ \forall t \in [a, b]$, existe entonces la función inversa $\psi^{-1}: [c, d] \to [a, b]$ la cual es también de clase $\mathscr{C}^1$ y además, como $(\psi^{-1})'(\psi(t))\psi'(t) = 1\ \forall t \in [a, b]$, también se tiene que $(\psi^{-1})'(s) \neq 0\ \forall s \in [c, d]$. Más aún, como $\psi'(t) > 0\ \forall t \in [a, b]$, también se tiene $(\psi^{-1})'(s) > 0\ \forall s \in [c, d]$.

Obsérvese que la función $\varphi = \psi^{-1}: [c, d] \to [a, b]$ tiene entonces todas las características que se necesitan para que $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ sea una reparametrización de $\mathbf{f}$. Estudiémosla.

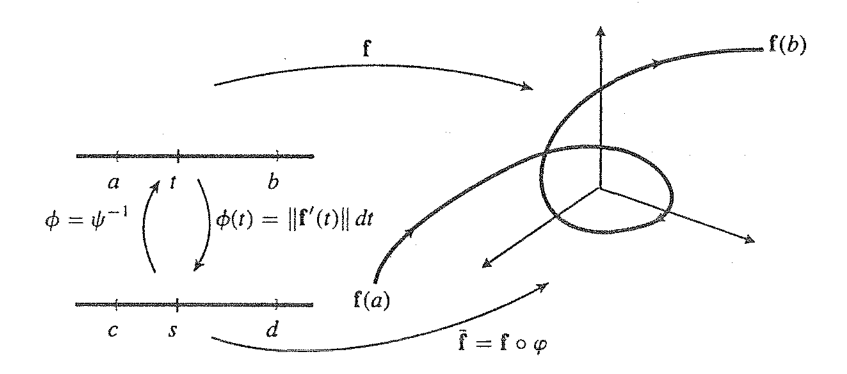


Figura 1. Gráfica de la curva C descrita por $\mathbf{f}$ y $\bar{\mathbf{f}}$.

Para $s \in [c, d]$ tenemos

$$\bar{\mathbf{f}}(s) = (\mathbf{f} \circ \varphi)(s) = \mathbf{f}(\varphi(s))$$

(Podemos poner a s como una longitud de una parte de la curva descrita por $\mathbf{f}$). Obsérvese que el vector velocidad $\bar{\mathbf{f}}'(s)$ es

$$\bar{\mathbf{f}}'(s) = \varphi'(s)\mathbf{f}'(\varphi(s))$$

Pero

$$\varphi'(s) = (\psi^{-1})'(s) = (\psi^{-1})'(\psi(t)) = \frac{1}{\psi'(t)} = \frac{1}{\|\mathbf{f}'(t)\|}$$

donde $s = \psi(t) = \int_{t_0}^{t} \|\mathbf{f}'(u)\|\, du$. Entonces

$$\|\bar{\mathbf{f}}(s)\| = \frac{1}{\|\mathbf{f}'(t)\|}\|\mathbf{f}'(t)\| = 1$$

De modo que $\bar{\mathbf{f}}$ es una reparametrización de $\mathbf{f}$ cuya propiedad es recorrer la curva de sus imágenes a una rapidez constante igual a la unidad. Diremos que $\bar{\mathbf{f}}$ es una *reparametrización de* $\mathbf{f}$ *por longitud de arco,* en el sentido de que la nueva variable independiente s de $\bar{\mathbf{f}}$ es justamente la longitud del camino entre t_0 y t,

$$s = \psi(t) = \int_{t_0}^{t} \|\mathbf{f}'(u)\|\, du$$

Nótese que el punto t_0 tomado del intervalo $[a, b]$ no jugó papel alguno en la discusión anterior. Podemos entonces, sin pérdida de generalidad, tomar siempre $t_0 = a$.

Debemos señalar que la reparametrización de un camino $\mathbf{f}$ por longitud de arco es una construcción de gran importancia teórica, pues a través de ella vamos a establecer nuevos conceptos que estudiaremos en la próximas secciones relacionados con la geométria de la curva descrita por $\mathbf{f}$. Los cálculos técnicos involucrados al tratar de llevar a cabo una de estas reparametrizaciones con algún camino dado, pueden ser sumamente complicados, o, en ocasiones, con impedimentos algebraicos que no permiten hacer explícita la función φ con la cual construimos la reparametrización (que es la función inversa de $s = \psi(s) = \int_a^t \|\mathbf{f}'(u)\|\,du$). Tendremos, sin embargo, algunos ejemplos tomados de situaciones que permiten ver, sin muchas dificultades, cómo se efectúan estas reparametrizaciones.

Ejemplo 1. Sea $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (r\cos t, r\operatorname{sen} t)$. Este es un camino regular que recorre una vez el círculo $x^2 + y^2 = r^2$. Obtengamos la reparametrización de $\mathbf{f}$ por longitud de arco. Se tiene

$$s = \psi(t) = \int_0^t \|\mathbf{f}'(u)\|\,du = \int_0^t \|(-r\operatorname{sen} u, r\cos u)\|\,du = \int_0^t r\,du = rt$$

de modo que la función $t = \varphi(s)$ es

$$t = \varphi(s) = \psi^{-1}(s) = \frac{1}{r}s$$

Entonces el camino $\bar{\mathbf{f}}: [0, 2\pi r] \to \mathbb{R}^2$, dado por

$$\bar{\mathbf{f}}(s) = (\mathbf{f}\circ\varphi)(s) = \mathbf{f}\left(\frac{1}{r}s\right) = \left(r\cos\frac{s}{r}, r\operatorname{sen}\frac{s}{r}\right)$$

es la reparametrización por longitud de arco de $\mathbf{f}$. Observe que se tiene

$$\|\bar{\mathbf{f}}'(s)\| = \left\|\left(-r\left(\operatorname{sen}\frac{s}{r}\right)\left(\frac{1}{r}\right), r\left(\cos\frac{s}{r}\right)\left(\frac{1}{r}\right)\right)\right\| = \left\|\left(-\operatorname{sen}\frac{s}{r}, \cos\frac{s}{r}\right)\right\| = 1$$

para todo $s \in [0, 2\pi r]$, como tenía que ocurrir. ■

Ejemplo 2. Sea $\mathbf{f}: [0, 3] \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (t, \cosh t)$. Se trata de un camino regular que recorre la catenaria $y = \cosh x$ del punto $(0, 1)$ al punto $(3, \cosh 3)$ con una velocidad dada por el vector

$$\mathbf{f}'(t) = (1, \operatorname{senh} t) \quad t \in [0, 3]$$

y una rapidez

$$\|\mathbf{f}'(t)\| = \sqrt{1 + \operatorname{senh}^2 t} = \cosh t, \quad t \in [0, 3]$$

Es decir, la velocidad con que se mueve el punto $\mathbf{f}(x)$ en la catenaria $y = \cosh x$ en el instante $t = x$ es justamente el valor de la ordenada de la curva correspondiente a esta abscisa. Obtengamos la reparametrización de $\mathbf{f}$ por longitud de arco, la cual recorrerá el mismo arco de catenaria con una rapidez constante igual a uno.

Tenemos

$$s = \psi(t) = \int_0^t \|\mathbf{f}'(u)\|\,du = \operatorname{senh} t - \operatorname{senh} 0 = \operatorname{senh} t$$

(ver ejemplo 8 de la sección anterior). Así que

$$t = \varphi(s) = \psi^{-1}(s) = \operatorname{arcsenh} s = \ln\left(s + \sqrt{s^2 + 1}\right)$$

La longitud de $\mathbf{f}$ entre $t = 0$ y $t = 3$ es senh 3 de modo que la reparametrización buscada es $\bar{\mathbf{f}}: [0, \operatorname{senh} 3] \to \mathbb{R}^2$ dada por

$$\bar{\mathbf{f}}(s) = \mathbf{f}(\varphi(s)) = \left(\ln(s + \sqrt{s^2+1}), \cosh(\ln(s + \sqrt{s^2+1}))\right)$$

Se tiene

$$\|\bar{\mathbf{f}}'(s)\|^2 = \left(\frac{1}{\sqrt{s^2+1}}, \frac{1}{\sqrt{s^2+1}} \operatorname{senh}(\ln(s + \sqrt{s^2+1}))\right)$$

de modo que

$$\begin{aligned} \|\bar{\mathbf{f}}'(s)\| &= \left(\frac{1}{\sqrt{s^2+1}}\right)^2 + \left(\frac{1}{\sqrt{s^2+1}} \operatorname{senh}(\ln(s + \sqrt{s^2+1}))\right)^2 \\ &= \frac{1}{s^2+1}\left(1 + \operatorname{senh}^2(\operatorname{arcsenh} s)\right) = \frac{1}{s^2+1}(1+s^2) = 1 \end{aligned}$$

como tenía que ocurrir. ■

Ejemplo 3. Un ejemplo "parecido" al del ejemplo anterior es $\mathbf{f}: [0,3] \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^2+1)$, pues la curva que éste describe es el arco de parábola $y = x^2 + 1$ en $0 \le x \le 3$, cuyo aspecto tiene cierto parecido con el arco de catenaria $y = \cosh x$ en $0 \le x \le 3$. (En el siglo XVIII se planteó el problema de describir la ecuación de la curva que formaba una cadena homogénea que colgaba libremente de dos puntos fijos de ella [con la misma altura]. Esta curva NO ES UNA PARÁBOLA, como se llegó a pensar inicialmente, sino una catenaria, nombre dado desde ese tiempo a la curva $y = \cosh x$ que es la solución del problema. Es obvio, sin embargo, el parecido geométrico de estas curvas). El problema de reparametrizar por longitud de arco el camino $\mathbf{f}$ es en este caso sumamente más complicado que el anterior, no obstante el parecido mencionado de las curvas. Veamos

$$s = \psi(t) = \int_0^t \|\mathbf{f}'(u)\|\, du = \int_0^t \|(1, 2u)\|\, du = \int_0^t \sqrt{1+4u^2}\, du$$

(La integral $I = \int \sqrt{a^2+x^2}\, dx$ se puede resolver por la sustitución trigonométrica $x = a\tan u$, conduciéndonos así a la integral $a \int \sec^3 u\, du$. Un procedimiento más sencillo es aplicar la sustitución hiperbólica $x = a \operatorname{senh} u$. En este caso la integral I se convierte en $a^2 \int \cosh^2 u\, du$, la cual se resuelve usando la identidad hiperbólica $\cosh^2 u = \frac{1}{2}(1+\cosh 2u)$. Queda como $I = a^2 \int \frac{1}{2}(1+\cosh 2u)\, du = a^2\left(\frac{u}{2} + \frac{1}{4}\operatorname{senh} u\right) = \frac{a^2}{2}(u + \operatorname{senh} u \cosh u)$, expresión que se ve después de regresar a la variable x, y de hacer algunas simplificaciones, como $\int \sqrt{a^2+x^2}\, dx = \frac{a^2}{2}\ln(x + \sqrt{a^2+x^2}) + \frac{x}{2}\sqrt{a^2+x^2}$)

Regresando al cálculo de $s = \psi(t)$, nos queda

$$\begin{aligned} s = \psi(t) &= \frac{1}{4}\ln\left(u + \frac{1}{2}\sqrt{1+4u^2}\right) + \frac{u}{2}\sqrt{1+4u^2}\,\Bigg|_0^t \\ &= \frac{1}{4}\ln\left(t + \frac{1}{2}\sqrt{1+4t^2}\right) + \frac{t}{2}\sqrt{1+4t^2} - \frac{1}{4}\ln\left(\frac{1}{2}\right) \\ &= \frac{1}{4}\ln\left(2t + \sqrt{1+4t^2}\right) + \frac{t}{2}\sqrt{1+4t^2} \end{aligned}$$

Cuando $t = 3$ tenemos $\alpha = \psi(3) = \frac{1}{4}\ln(6+\sqrt{37})+\frac{3}{2}\sqrt{37}$. Ciertamente la función $\psi: [0, 3] \to [0, \alpha]$ es sobreyectiva, de clase $\mathscr{C}^1$, y $\psi'(t) = \sqrt{1+t^2} > 0\ \forall t \in [0, 3]$, de modo que podemos hablar de su inversa $\varphi = \psi^{-1}: [0, \alpha] \to [0, 3]$. Esta función $t = \varphi(s)$ es la que necesitamos para construir la reparametrización $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$ de $\mathbf{f}$ por longitud de arco. Sin embargo, es evidente que en este caso no es posible hacer explícita la función φ. Así pues, nos conformamos con decir que el camino que reparametriza a $\mathbf{f}$ por longitud de arco es $\bar{\mathbf{f}}: [0, \frac{1}{4}\ln(6+\sqrt{37}) + \frac{3}{2}\sqrt{37}] \to \mathbb{R}^2$, $\bar{\mathbf{f}}(s) = \mathbf{f}(\varphi(s)) = (\varphi(s), (\varphi(s))^2 + 1)$, donde $\varphi(s)$ es la inversa de la función

$$s = \psi(t) = \frac{1}{4}\ln(2t + \sqrt{1+4t^2}) + \frac{t}{2}\sqrt{1+4t^2} \qquad \blacksquare$$

Ejemplo 4. Sea $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ el camino $\mathbf{f}(t) = (\alpha\cos t, \alpha \operatorname{sen} t, \beta t)$. La curva que éste describe es una hélice circular (que se dibuja en el cilindro $x^2 + y^2 = \alpha^2$). Tomemos como punto base $\mathbf{f}(t_0)$ el correspondiente en $t_0 = 0$ (es decir, el punto $(\alpha, 0, 0)$) a partir del cual mediremos la longitud del camino. Se tiene

$$\begin{aligned} s = \psi(t) &= \int_0^t \|\mathbf{f}'(u)\|\,du = \int_0^t \|(-\alpha \operatorname{sen} u, \alpha\cos u, \beta)\|\,du \\ &= \int_0^t \sqrt{\alpha^2+\beta^2}du = \sqrt{\alpha^2+\beta^2}t \end{aligned}$$

de modo que

$$t = \varphi(s) = \psi^{-1}(s) = \frac{1}{\sqrt{\alpha^2+\beta^2}}s$$

La reparametrización de $\mathbf{f}$ por longitud de arco es entonces el camino $\bar{\mathbf{f}}: \mathbb{R} \to \mathbb{R}^3$ dado por

$$\bar{\mathbf{f}}'(s) = \mathbf{f}(\varphi(s)) = \left(\alpha\cos\frac{s}{\sqrt{\alpha^2+\beta^2}}, \alpha \operatorname{sen}\frac{s}{\sqrt{\alpha^2+\beta^2}}, \frac{\beta s}{\sqrt{\alpha^2+\beta^2}}\right) \qquad \blacksquare$$

Ejercicios (Capítulo 5, Sección 6)

1. Considere el camino $\mathbf{f}: [0, 2\pi r] \to \mathbb{R}^2$ dado por

$$\mathbf{f}(s) = \left(r\cos\frac{s}{r}, r \operatorname{sen}\frac{s}{r}\right)$$

Este es la parametrización por longitud de arco del círculo $x^2 + y^2 = r^2$ (ver ejemplo 1). Calcule $\mathbf{f}''(s)$. Compruebe que este vector es ortogonal a $\mathbf{f}'(s)$ para toda s en $[0, 2\pi r]$. Interprete geométricamente este hecho. Calcule $\|\mathbf{f}''(s)\|$.

2. Considere el camino $\mathbf{f}: [0, \operatorname{senh} 3] \to \mathbb{R}^2$ dado por

$$\mathbf{f}(s) = (\ln(s + \sqrt{s^2+1}), \cosh(\ln(s+\sqrt{s^2+1})))$$

Este es la parametrización por longitud de arco de la catenaria $y = \cosh x$ en el intervalo $[0, 3]$ (ver ejemplo 2). Calcule $\mathbf{f}''(s)$. Compruebe que este vector es ortogonal a $\mathbf{f}'(s)$ para toda s en $[0, \operatorname{senh} 3]$. Calcule $\|\mathbf{f}''(s)\|$.

3. Considere el camino $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^3$ dado por

$$\mathbf{f}(s) = \left(\alpha \cos \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \alpha \operatorname{sen} \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \frac{\beta s}{\sqrt{\alpha^2 + \beta^2}}\right)$$

Este es la parametrización por longitud de arco de la hélice circular $\mathbf{g}(t) = (\alpha \cos t, \alpha \operatorname{sen} t, \beta t)$, $t \in \mathbb{R}$ (ver ejemplo 4). Calcule $\mathbf{f}''(s)$. Constate que este vector es ortogonal a $\mathbf{f}'(s)$ para toda $s \in \mathbb{R}$. Calcule $\|\mathbf{f}''(s)\|$.

4. Sea $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino parametrizado por longitud de arco. Demuestre que el vector $\mathbf{f}''(s)$ es ortogonal al vector $\mathbf{f}'(s)$ $\forall s \in I$.

5.7 Curvatura

En esta sección estudiaremos uno de los dos conceptos más importantes en el estudio de las curvas, llamado curvatura; el otro, llamado torsión, será objeto de estudio de la sección 10. La idea general que se persigue en el estudio de la curvatura de una curva es la de medir la rapidez con que la curva se aleja de su recta tangente en un punto $\mathbf{p}$ dado de ella. En términos generales, a esta rapidez se le llama "curvatura de la curva en el punto $\mathbf{p}$". La idea intuitiva subyacente en este ejemplo tiene que ver con "qué tanto se curva la curva en el punto $\mathbf{p}$". Por ejemplo, al ver las siguientes figuras

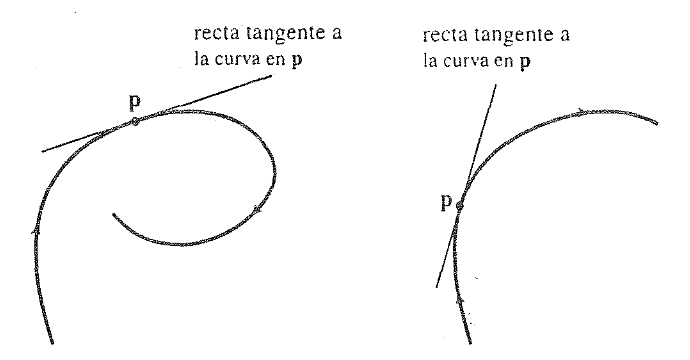


Figura 1. Dos curvas con diferente curvatura en el punto $\mathbf{p}$.

podemos intuir que la curvatura de la primera en $\mathbf{p}$ es más grande que la de la segunda en $\mathbf{p}$. Es claro entonces que el concepto de curvatura es un concepto *local*: hablamos de la curvatura de una curva *en un punto* $\mathbf{p}$ *de ella*. También debe ser claro que para iniciar este estudio, necesitamos trabajar con curvas que sean imágenes de caminos regulares, ya que a ellos son a los que siempre podemos asociar rectas tangentes. En realidad, tendremos que pedir un poco más a los caminos $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$ con los que trabajaremos en esta sección. Estos deben ser tales que su segunda derivada $\mathbf{f}''(t)$ deba existir para $t \in I$. Si $\mathbf{f}(t) = (x_1(t), \ldots, x_n(t))$, el vector segunda derivada de $\mathbf{f}$ es $\mathbf{f}''(t) = (x_1''(t), \ldots, x_n''(t))$. Sea entonces $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino regular para el que $\mathbf{f}''(t)$ existe, $\forall t \in I$ (diremos que $\mathbf{f}$ es un camino dos veces diferenciable en I). Hemos visto que $\|\mathbf{f}'(t)\|$ nos da información de la rapidez con que el punto $\mathbf{f}(t)$ se está moviendo sobre la curva de las imágenes del camino $\mathbf{f}$. Obsérvese entonces que el número positivo $\|\mathbf{f}'(t)\|$ nos da la *rapidez de*

variación de las imágenes del camino $\mathbf{f}$ *en el punto* $\mathbf{f}(t)$. Con esta perspectiva en mente, el número $\|\mathbf{f}''(t)\|$ nos daría información sobre la rapidez de variación del vector velocidad $\mathbf{f}'(t)$ en el punto $\mathbf{f}(t)$. Al hablar de variación de un vector, estamos incluyendo en ésta la variación de la magnitud del vector y la variación de la dirección del vector. Por ejemplo el camino $\mathbf{f}: \mathbb{R}^+ \to \mathbb{R}^2$, dado por $\mathbf{f}(t) = (t^3, t^3)$ cuyas imágenes describen la recta $y = x$ en el primer cuadrante, tiene por vector velocidad a $\mathbf{f}'(t) = (3t^2, 3t^2)$; ciertamente este vector tiene siempre la misma dirección, pero su magnitud es variable, de modo que $\|\mathbf{f}''(t)\| = \|(6t, 6t)\| = 6\sqrt{2}t$ nos da solamente la información de la variación de la magnitud del vector $\mathbf{f}'(t)$ en el punto $\mathbf{f}(t)$. Por otro lado, el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, dado por $\mathbf{f}(t) = (\cos 2t, \operatorname{sen} 2t)$ tiene por vector velocidad a $\mathbf{f}'(t) = (-2 \operatorname{sen} 2t, 2\cos 2t)$. Nótese que $\|\mathbf{f}'(t)\| = 2\ \forall t \in \mathbb{R}$, de modo que en este caso la magnitud del vector velocidad $\mathbf{f}'(t)$ es constante, no así su dirección. Entonces el número $\|\mathbf{f}''(t)\| = \|(-4\cos 2t, -4\operatorname{sen} 2t)\| = 4$ nos da información de la variación de la dirección del vector $\mathbf{f}'(t)$ en el punto $\mathbf{f}(t)$.

Al estudiar la curvatura de una curva en un punto $\mathbf{p}$ de ella nos interesará tener información sobre la variación del vector tangente $\mathbf{f}'(t)$ en el punto $\mathbf{p}$. Por supuesto que lo único que nos debe interesar es la *variación en la dirección* de $\mathbf{f}'(t)$, y no en su magnitud, de tal manera que, por ejemplo, para el camino $\mathbf{f}(t) = (t^3, t^3)$ mencionado anteriormente, cuyo vector velocidad no cambia de dirección, podamos decir que "su curvatura es cero", coincidiendo así con la idea intuitiva que sembramos al comienzo de la sección (¡una recta no se curva!).

Con estas consideraciones previas, resulta entonces claro que para tener una medida de la rapidez de variación en la dirección del vector velocidad de un camino, debemos tomar un camino que tenga rapidez $\|\mathbf{f}'(t)\|$ constante, digamos igual a la unidad, pues en tal caso $\|\mathbf{f}''(t)\|$ sería una medida (¡la medida que necesitamos!) de cuánto está cambiando *la dirección* del vector $\mathbf{f}'(t)$ de la curva en el punto $\mathbf{f}(t)$.

Luego de este preámbulo, no debe resultar extraño que el estudio emprendido en la sección anterior fue precisamente para resolver esta dificultad técnica que ahora se nos presenta; para introducir el concepto de curvatura lo haremos por medio de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ reparametrizado por longitud de arco, cuya característica fundamental es precisamente que su rapidez $\|\mathbf{f}'(t)\|$ es constante e igual a 1 para todo $t \in I$. Establezcamos rigurosamente entonces el concepto de curvatura.

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ un camino dos veces diferenciable parametrizado por longitud de arco.[1] Al número $k(s) = \|\mathbf{f}''(s)\|$ se le llama *curvatura* de $\mathbf{f}$ en s. ■

Seguimos respetando el uso de la letra s para denotar a la variable independiente (la longitud de arco) de un camino reparametrizado por longitud de arco, para el cual, además, usaremos la notación $\mathbf{T}(s)$ para designar al vector $\mathbf{f}'(s)$, llamado *vector tangente unitario* de $\mathbf{f}$ en s. Entonces, $\mathbf{T}(s)$ es el vector tangente (el vector velocidad) del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^n$ parametrizado por longitud de arco, y así, $\|\mathbf{T}(s)\| = 1\ \forall s \in I$. Con esta notación, la curvatura de $\mathbf{f}$ en s se ve como $k(s) = \|\mathbf{T}'(s)\|$.

Ejemplo 1. (La curvatura de una recta es cero). Una recta en $\mathbb{R}^3$ que pasa por el punto $\mathbf{p} = (x_0, y_0, z_0)$ y tiene a $\mathbf{v} = (a, b, c)$ por vector paralelo, se puede ver como la imagen del camino $\mathbf{f}^*: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}^*(t) = (x_0, y_0, z_0) + t(a, b, c)$. Es fácil ver que la reparametrización de $\mathbf{f}^*$ por longitud

[1] Este camino $\mathbf{f}$ al que se refiere la definición, es en realidad el camino $\tilde{\mathbf{f}}$ que manejamos en la sección anterior, reparametrización de un camino por longitud de arco. Por razones de simplicidad en la notación, pensemos que el camino dado $\mathbf{f}$ ya es la reparametrización por longitud de arco de un camino previamente dado $\mathbf{f}^*$. En tal caso decimos que $\mathbf{f}$ está "parametrizado" por longitud de arco.

de arco es el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por

$$\mathbf{f}(s) = (x_0, y_0, z_0) + \frac{s}{\sqrt{a^2 + b^2 + c^2}}(a, b, c) = \mathbf{p} + \mathbf{u}s$$

donde $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$. Observe entonces que $\mathbf{f}'(s) = \mathbf{u}$ y $\mathbf{f}''(s) = \mathbf{0}$ (el vector $\mathbf{0}$ de $\mathbb{R}^3$), de modo que $k(s) = \|\mathbf{f}''(s)\| = 0$, $\forall s \in \mathbb{R}$. Es decir, la curvatura de $\mathbf{f}$ en cualquier punto es 0. ■

Ejemplo 2. (La curvatura de un círculo es constante). Sea $\mathbf{f}: [0, 2\pi r] \to \mathbb{R}^2$ el camino $\mathbf{f}(s) = \left(r \cos \frac{s}{r}, r \operatorname{sen} \frac{s}{r}\right)$. Esta es la parametrización por longitud de arco del círculo $x^2 + y^2 = r^2$ (ver ejemplo 1 de la sección anterior). Tenemos

$$\mathbf{T}(s) = \mathbf{f}'(s) = \left(\operatorname{sen} \frac{s}{r}, \cos \frac{s}{r}\right)$$

$$\mathbf{f}''(s) = \left(-\frac{1}{r} \cos \frac{s}{r}, -\frac{1}{r} \operatorname{sen} \frac{s}{r}\right)$$

de modo que la curvatura de $\mathbf{f}$ en s es

$$k(s) = \|\mathbf{f}''(s)\| = \frac{1}{r} \qquad \forall s \in [0, 2\pi]$$

Podemos decir entonces que la curvatura de un círculo de radio r es constante e igual a $\frac{1}{r}$. Esto responde a la idea intuitiva de que en un círculo el vector tangente unitario $\mathbf{T}(s)$ tiene la misma rapidez de variación en su dirección en todos los puntos, o bien, que la curva se aleja de su tangente a la misma rapidez en todos sus puntos. ■

Ejemplo 3. Consideremos el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por

$$\mathbf{f}(s) = \left(\alpha \cos \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \alpha \operatorname{sen} \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \frac{\beta s}{\sqrt{\alpha^2 + \beta^2}}\right).$$

Esta es la parametrización por longitud de arco de una hélice en $\mathbb{R}^3$ (ver ejemplo 4 de la sección anterior). Tenemos

$$\mathbf{T}(s) = \mathbf{f}'(s) = \left(\frac{-\alpha}{\sqrt{\alpha^2 + \beta^2}} \operatorname{sen} \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} \cos \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}\right)$$

La segunda derivada es

$$\mathbf{f}''(s) = \left(\frac{-\alpha}{\alpha^2 + \beta^2} \cos \frac{s}{\sqrt{\alpha^2 + \beta^2}}, \frac{-\alpha}{\alpha^2 + \beta^2} \operatorname{sen} \frac{s}{\sqrt{\alpha^2 + \beta^2}}, 0\right)$$

de modo que la curvatura de $\mathbf{f}$ en s es

$$k(s) = \|\mathbf{f}''(s)\| = \frac{|\alpha|}{\alpha^2 + \beta^2}$$

Obsérvese que al igual que sucedió en el ejemplo anterior, la curvatura es constante en todos los puntos de la curva. ■

Si bien es cierto que la definición que dimos de curvatura muestra con toda claridad y sencillez la esencia del concepto, también es cierto que tal definición resulta impráctica si quisiéramos (como de hecho queremos) hacer cálculos concretos de curvaturas de curvas, pues para poder hacerlo tendríamos que pasar antes por la reparametrización por longitud de arco del camino. En otras palabras, el proceso que sugiere la definición de curvatura para hacer cálculos con ella es la siguiente: dado el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ o $\mathbb{R}^3$ del cual se quiere calcular su curvatura en cierto punto correspondiente a $t_0 \in I$, obtenga primero la reparametrización por longitud de arco $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^2$ o $\mathbb{R}^3$; con ella, calcule $\bar{\mathbf{f}}''(s)$; la norma de este vector evaluado en $s_0 = \psi^{-1}(t_0)$, donde $\psi(t) = \int_0^t \|\mathbf{f}'(u)\|\, du$ es la curvatura buscada. Además de que ciertamente resulta molesto pensar en que para cada camino hay que pasar primero por su reparametrización por longitud de arco, acontece que más allá de la incomodidad que esto representa se dan curvas tan simples como $y = x^2$, imagen del camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^2)$, para los que *no* o podríamos calcular, con la definición dada, su curvatura (ver ejemplo 3 de la sección anterior).

Veamos cómo podemos hacer los cálculos de la curvatura de un camino sin pasar por la reparametrización por longitud de arco; es decir, veamos cómo expresar la segunda derivada $\bar{\mathbf{f}}''(s)$ de la reparametrización por longitud de arco de un camino dado $\mathbf{f}(t)$ en términos de la o las derivadas de éste. Haremos este análisis en el caso concreto de caminos en $\mathbb{R}^3$.

Sea entonces $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ un camino regular dado, dos veces diferenciable y sea $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^3$ su reparametrización por longitud de arco. Recuerde el esquema

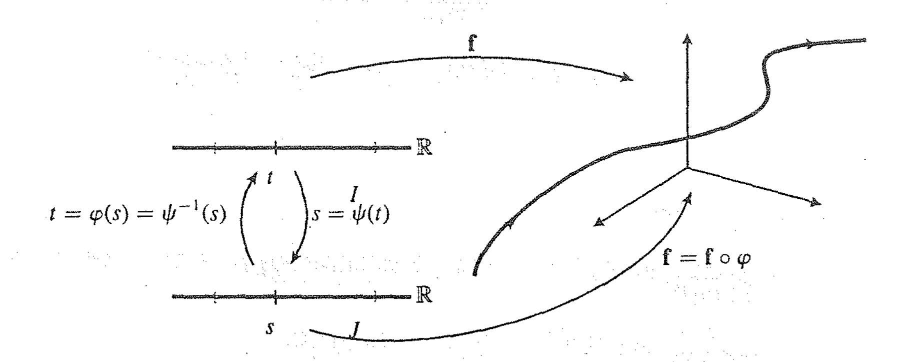


Figura 2. Reparametrización de un camino por longitud de arco.

Lo que queremos hacer es dejar expresada la norma del vector $\bar{\mathbf{f}}''(s)$ (la curvatura) en términos de la norma de los vectores $\mathbf{f}'(t)$ y $\mathbf{f}''(t)$ (que se calculan directamente del camino dado).

De la misma definición de $\bar{\mathbf{f}}'(s) = \mathbf{T}(s)$ es claro que

$$\bar{\mathbf{f}}'(s) = \frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t) \text{ en donde } t = \varphi(s) = \psi^{-1}(s)$$

o bien, haciendo directamente los cálculos en $\bar{\mathbf{f}}(s) = \mathbf{f}(\varphi(s))$, obtenemos, derivando $\bar{\mathbf{f}}'(s) = \varphi'(s)\mathbf{f}'(\varphi(s))$. Pero $\varphi'(s) = (\psi^{-1})'(s) = \frac{1}{\psi'(\varphi(s))} = \frac{1}{\psi'(t)} = \frac{1}{\|\mathbf{f}'(t)\|}$, de modo que

$$\bar{\mathbf{f}}'(s) = \frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t)$$

Derivando nuevamente respecto de s obtenemos

$$\begin{aligned}\bar{\mathbf{f}}''(s) &= \frac{d}{ds}\left(\frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t)\right) = \frac{d}{dt}\left(\frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t)\right)\frac{dt}{ds} \\ &= \frac{d}{dt}\left(\frac{\mathbf{f}'(t)}{\|\mathbf{f}'(t)\|}\right)\varphi'(s) \\ &= \left(\frac{\|\mathbf{f}'(t)\|\frac{d}{dt}\mathbf{f}'(t) - \mathbf{f}'(t)\frac{d}{dt}\|\mathbf{f}'(t)\|}{\|\mathbf{f}'(t)\|^2}\right)\left(\frac{1}{\|\mathbf{f}'(t)\|}\right)\end{aligned}$$

Recordando que

$$\frac{d}{dt}\|\mathbf{g}(t)\| = \frac{\mathbf{g}(t)\cdot\mathbf{g}'(t)}{\|\mathbf{g}(t)\|}$$

(ver página 453), nos queda

$$\begin{aligned}\bar{\mathbf{f}}''(s) &= \frac{\|\mathbf{f}'(t)\|\mathbf{f}''(t) - \mathbf{f}'(t)\dfrac{\mathbf{f}'(t)\cdot\mathbf{f}''(t)}{\|\mathbf{f}'(t)\|}}{\|\mathbf{f}'(t)\|^3} \\ &= \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\right)\end{aligned}$$

Entonces

$$\begin{aligned}\|\bar{\mathbf{f}}''(s)\|^2 &= \bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}''(s) \\ &= \frac{1}{\|\mathbf{f}'(t)\|^8}\left[\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - \left(\mathbf{f}'(t)\cdot\mathbf{f}''(t)\right)\mathbf{f}'(t)\right]\cdot\left[\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - \left(\mathbf{f}'(t)\cdot\mathbf{f}''(t)\right)\mathbf{f}'(t)\right] \\ &= \frac{1}{\|\mathbf{f}'(t)\|^8}\left(\|\mathbf{f}'(t)\|^4\mathbf{f}''(t)\cdot\mathbf{f}''(t) - 2\|\mathbf{f}'(t)\|^2\left(\mathbf{f}'(t)\cdot\mathbf{f}''(t)\right)^2\right. \\ &\qquad \left. + \left(\mathbf{f}'(t)\cdot\mathbf{f}''(t)\right)^2\mathbf{f}'(t)\cdot\mathbf{f}'(t)\right) \\ &= \frac{1}{\|\mathbf{f}'(t)\|^8}\left(\|\mathbf{f}'(t)\|^4\|\mathbf{f}''(t)\|^2 - \|\mathbf{f}'(t)\|^2\left(\mathbf{f}'(t)\cdot\mathbf{f}''(t)\right)^2\right) \\ &= \frac{1}{\|\mathbf{f}'(t)\|^6}\left(\|\mathbf{f}'(t)\|^2\|\mathbf{f}''(t)\|^2 - \left(\mathbf{f}'(t)\cdot\mathbf{f}''(t)\right)^2\right)\end{aligned}$$

Recuerde que si $\mathbf{u}$, $\mathbf{v}$ son dos vectores de $\mathbb{R}^3$ entonces

$$\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u}\cdot\mathbf{v})^2$$

Con esta fórmula podemos escribir entonces

$$\|\bar{\mathbf{f}}''(s)\|^2 = \frac{1}{\|\mathbf{f}'(t)\|^6}\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|^2$$

de donde

$$\|\bar{\mathbf{f}}''(s)\| = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3}$$

Así pues, la curvatura del camino regular $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ dos veces diferenciable, que denotaremos por $k(t)$, es

$$\boxed{k(t) = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3}}$$

Esta es la fórmula que buscábamos, con la cual podremos calcular la curvatura del camino $\mathbf{f}$ trabajando directamente con él.

Ejemplo 4. El camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por $\mathbf{f}(t) = (\alpha \cos t, \alpha \operatorname{sen} t, \beta t)$ es el mismo que el considerado en el ejemplo 3, sólo que estaba parametrizado por longitud de arco. Tenemos

$$\begin{aligned} \mathbf{f}'(t) \times \mathbf{f}''(t) &= (-\alpha \operatorname{sen} t, \alpha \cos t, \beta) \times (-\alpha \cos t, -\alpha \operatorname{sen} t, 0) \\ &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\alpha \operatorname{sen} t & \alpha \cos t & \beta \\ -\alpha \cos t & -\alpha \operatorname{sen} t & 0 \end{bmatrix} = (\alpha\beta \operatorname{sen} t, -\alpha\beta \cos t, \alpha^2) \end{aligned}$$

de modo que

$$\|\mathbf{f}'(t) \times \mathbf{f}''(t)\| = \sqrt{\alpha^2\beta^2 \operatorname{sen}^2 t + \alpha^2\beta^2 \cos^2 t + \alpha^4} = |\alpha|\sqrt{\alpha^2 + \beta^2}$$

y como

$$\|\mathbf{f}'(t)\| = \|(-\alpha \operatorname{sen} t, \alpha \cos t, \beta)\| = \sqrt{\alpha^2 + \beta^2}$$

tenemos que la curvatura $k(t)$ es

$$k(t) = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3} = \frac{|\alpha|\sqrt{\alpha^2 + \beta^2}}{\left(\sqrt{\alpha^2 + \beta^2}\right)^3} = \frac{|\alpha|}{\alpha^2 + \beta^2}$$

resultado que coincide con el obtenido en el ejemplo 3. ■

Ejemplo 5. Considere el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por $\mathbf{f}(t) = (2 + t, 1 + t^2, 3t + t^2)$. Calculemos su curvatura $k(t)$. Se tiene $\mathbf{f}'(t) = (1, 2t, 3 + 2t)$, $\mathbf{f}''(t) = (0, 2, 2)$, de modo que

$$\mathbf{f}'(t) \times \mathbf{f}''(t) = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2t & 3 + 2t \\ 0 & 2 & 2 \end{bmatrix} = (-6, -2, 2)$$

Entonces

$$k(t) = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3} = \frac{\|(-6, -2, 2)\|}{\|(1, 2t, 3 + 3t)\|^3} = \frac{\sqrt{44}}{(10 + 12t + 8t^2)^{3/2}}$$

Observe que en este caso la curvatura $k(t)$ sí depende de t. Por ejemplo, $k(0) = \frac{\sqrt{44}}{(10)^{3/2}}$, $k(1) = \frac{\sqrt{44}}{(30)^{3/2}} < \frac{\sqrt{44}}{(10)^{3/2}} k(0)$, así que podemos decir que el camino $\mathbf{f}$, o bien, que la curva en $\mathbb{R}^3$ que este camino describe, se curva más en el punto $\mathbf{f}(0) = (2, 1, 0)$ que en el punto $\mathbf{f}(1) = (3, 2, 4)$. Más aún, considerando la "función curvatura", $k: \mathbb{R} \to \mathbb{R}$, vemos que

$$k'(t) = -\frac{3}{2}\sqrt{44}(10 + 12t + 8t^2)^{-5/2}(12 + 16t)$$

donde $k'(t) = 0 \Leftrightarrow t = -\frac{3}{4}$. En este valor de t se localiza un punto crítico para $k(t)$. Es fácil ver que se trata de un punto de máximo local. Entonces concluimos que el camino $\mathbf{f}$ tiene curvatura máxima cuando $t = -\frac{3}{4}$, es decir, en el punto $\mathbf{f}\left(-\frac{3}{4}\right) = \left(\frac{5}{4}, \frac{25}{16}, -\frac{27}{16}\right)$, la cual es $k\left(-\frac{3}{4}\right) = \sqrt{\frac{352}{1331}}$. ■

Quisiéramos ahora ver que la fórmula obtenida para calcular la curvatura de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ (regular, dos veces diferenciable) es invariante por reparametrizaciones del camino. Este hecho nos permite ver a la curvatura como una propiedad geométrica de la curva en sí, imagen del camino $\mathbf{f}$, independientemente de "la manera" como es recorrida.

En efecto, sea $\mathbf{g}: J \subseteq \mathbb{R} \to \mathbb{R}^3$, un camino regular dos veces diferenciable, reparametrización del camino $\mathbf{f}$. Entonces $\mathbf{g} = \mathbf{f} \circ \varphi$, con $\varphi: I \to J$ la inversa de la función $\psi: J \to I$

$$s = \psi(t) = \int_{t_0}^{t} \|\mathbf{f}'(u)\|\, du$$

Queremos ver que

$$k(s) = \frac{\|\mathbf{g}'(s) \times \mathbf{g}''(s)\|}{\|\mathbf{g}'(s)\|^3} = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3} = k(t)$$

donde $t = \varphi(s)$. Tenemos

$$\begin{aligned}
\mathbf{g}'(s) &= \varphi'(s)\mathbf{f}'(\varphi(s)) = \varphi'(s)\mathbf{f}'(t) \\
\mathbf{g}''(s) &= (\varphi'(s))^2\mathbf{f}''(\varphi(s)) + \varphi''(s)\mathbf{f}'(\varphi(s)) \\
&= (\varphi'(s))^2\mathbf{f}''(t) + \varphi''(s)\mathbf{f}'(t)
\end{aligned}$$

Entonces

$$\begin{aligned}
\mathbf{g}'(s) \times \mathbf{g}''(s) &= \varphi'(s)\mathbf{f}'(t) \times \left[(\varphi'(s))^2\mathbf{f}''(t) + \varphi''(s)\mathbf{f}'(t)\right] \\
&= (\varphi'(s))^3\mathbf{f}'(t) \times \mathbf{f}''(t) + \varphi'(s)\varphi''(s)\mathbf{f}'(t) \times \mathbf{f}'(t) \\
&= (\varphi'(s))^3\mathbf{f}'(t) \times \mathbf{f}''(t)
\end{aligned}$$

de modo que

$$\begin{aligned}
k(s) &= \frac{\|\mathbf{g}'(t) \times \mathbf{g}''(t)\|}{\|\mathbf{g}'(t)\|^3} = \frac{\|(\varphi'(s))^3\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\varphi'(s)\mathbf{f}'(t)\|^3} \\
&= \frac{|\varphi'(s)|^3\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{|\varphi'(s)|^3\|\mathbf{f}'(t)\|^3} = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3} = k(t), \qquad t = \varphi(s)
\end{aligned}$$

como queríamos comprobar.

Ejemplo 6. Consideremos el cilindro elíptico $2x^2 + 3y^2 = 1$ y el plano $z = 2y$. Estas dos superficies se intersectan en una elipse C. Calcularemos la curvatura de ésta en el punto $\mathbf{p} = \left(0, \frac{1}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right)$. Según lo expuesto anteriormente, podemos tomar cualquier camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ regular dos veces diferenciable cuya imagen sea la elipse C (al menos en los alrededores del punto $\mathbf{p}$) y con él calcular la curvatura deseada. Por ejemplo, el camino $\mathbf{f}: [0, 2\pi] \to \mathbb{R}^3$ dado por

$$\mathbf{f}(t) = \left(\frac{1}{\sqrt{2}}\cos t, \frac{1}{\sqrt{3}}\operatorname{sen} t, \frac{2}{\sqrt{3}}\operatorname{sen} t\right)$$

El punto $\mathbf{p}$ corresponde a $\mathbf{f}\left(\frac{\pi}{2}\right)$. Tenemos

$$\mathbf{f}'(t) = \left(-\frac{1}{\sqrt{2}}\operatorname{sen} t, \frac{1}{\sqrt{3}}\cos t, \frac{2}{\sqrt{3}}\cos t\right)$$

$$\mathbf{f}''(t) = \left(-\frac{1}{\sqrt{2}}\cos t, -\frac{1}{\sqrt{3}}\operatorname{sen} t, -\frac{2}{\sqrt{3}}\operatorname{sen} t\right)$$

En $t = \frac{\pi}{2}$ estos vectores son $\mathbf{f}'\left(\frac{\pi}{2}\right) = \left(-\frac{1}{\sqrt{2}}, 0, 0\right)$, $\mathbf{f}''\left(\frac{\pi}{2}\right) = \left(0, -\frac{1}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right)$. Entonces

$$\mathbf{f}'\left(\frac{\pi}{2}\right) \times \mathbf{f}''\left(\frac{\pi}{2}\right) = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\frac{1}{\sqrt{2}}, & 0 & 0 \\ 0 & -\frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} \end{bmatrix}$$
$$= \left(0, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

de modo que la curvatura de la elipse C en $\mathbf{p}$ es

$$k\left(\frac{\pi}{2}\right) = \frac{\left\|\mathbf{f}'\left(\frac{\pi}{2}\right) \times \mathbf{f}''\left(\frac{\pi}{2}\right)\right\|}{\left\|\mathbf{f}'\left(\frac{\pi}{2}\right)\right\|^3} = \frac{\left\|\left(0, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)\right\|}{\left\|\left(-\frac{1}{\sqrt{2}}, 0, 0\right)\right\|^3} = \frac{\left(\frac{5}{6}\right)^{1/2}}{\left(\frac{1}{2}\right)^{3/2}} = 2\sqrt{\frac{5}{3}}$$ ■

La discusión sobre curvatura de curvas planas, imágenes de caminos $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, se obtiene fácilmente como un caso particular de la presentada previamente para caminos en el espacio $\mathbb{R}^3$. En efecto, una curva plana, imagen del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (x(t), y(t))$ puede ser considerada como imagen del camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (x(t), y(t), 0)$, la cual queda dibujada en el plano xy $(z = 0)$.

Sea entonces $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ un camino regular dos veces diferenciable. Calculemos su curvatura $k(t)$. Tenemos

$$\mathbf{f}'(t) = (x'(t), y'(t), 0) \qquad \mathbf{f}''(t) = (x''(t), y''(t), 0)$$

y entonces

$$\mathbf{f}'(t) \times \mathbf{f}''(t) = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x'(t) & y'(t) & 0 \\ x''(t) & y''(t) & 0 \end{bmatrix} = (0, 0, x'(t)y''(t) - x''(t)y'(t))$$

de modo que

$$k(t) = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)^3\|} = \frac{|x'(t)y''(t) - x''(t)y'(t)|}{((x'(t))^2 + (y'(t))^2)^{3/2}}$$

esta la fórmula buscada que nos da la curvatura $k(t)$ del camino $\mathbf{f}$. Llamemos la atención al hecho de que la curvatura de un camino tal como ha sido definida, es siempre un número no negativo; sin embargo, en el caso de caminos en $\mathbb{R}^2$, es posible asociar un signo a la curvatura, el cual nos ampliará la información geométrica de la curva que el camino describe. Para ver esto retomaremos la discusión desde la definición misma de curvatura, para el caso de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (x(t), y(t))$ —regular dos veces diferenciable—.

Si $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^2$ es la reparametrización de $\mathbf{f}$ por longitud de arco, el vector $\mathbf{T}(s) = \bar{\mathbf{f}}'(s)$ es un vector unitario para toda $s \in J$. De $\|\bar{\mathbf{f}}'(s)\| = 1$, obtenemos, derivando respecto de s, que

$$0 = \frac{d}{ds}\|\mathbf{T}(s)\| = \frac{\mathbf{T}(s) \cdot \mathbf{T}'(s)}{\|\mathbf{T}(s)\|} = \mathbf{T}(s) \cdot \mathbf{T}'(s)$$

por lo que el vector $\mathbf{T}'(s) = \bar{\mathbf{f}}''(s)$ es un vector ortogonal a $\mathbf{T}(s)$ para toda $s \in J$. Por otra parte, consideremos el vector unitario $\mathbf{N}(s)$ que se obtiene al girar el vector $\mathbf{T}(s)$ un ángulo de $\frac{\pi}{2}$ en sentido contrario de las manecillas del reloj. Tenemos así que los vectores $\mathbf{T}'(s)$ y $\mathbf{N}(s)$ son ortogonales a $\mathbf{T}(s)$, y por lo tanto colineales, de modo que para cada $s \in J$, existe un número $k(s)$ bien definido

$$\mathbf{T}'(s) = \bar{\mathbf{f}}''(s) = k(s)\mathbf{N}(s)$$

a este número $k(s)$ lo llamaremos curvatura de $\bar{\mathbf{f}}$ en s.

Nótese que se trata de la misma definición que la dada anteriomente para un camino en $\mathbb{R}^3$, excepto, posiblemente, por un cambio de signo, pues

$$\|\bar{\mathbf{f}}''(s)\| = \|k(s)\mathbf{N}(s)\| = |k(s)|\|\mathbf{N}(s)\| = |k(s)|$$

El único detalle adicional que tiene nuestra nueva definición, es, entonces que ahora la curvatura puede ser positiva, cero, o *negativa*.

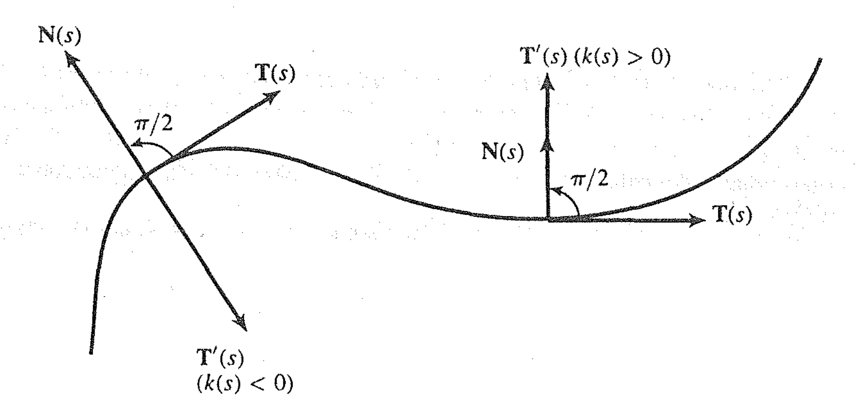


Figura 3. Los vectores N(s) y T(s) en una curva plana.

Obtengamos una expresión para la curvatura así definida. Tenemos que

$$\mathbf{T}(s) = \frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t) = \frac{1}{\|\mathbf{f}'(t)\|}(x'(t), y'(t)), \quad t = \varphi(s)$$

de modo que el vector $\mathbf{N}(s)$ debe ser

$$\mathbf{N}(s) = \frac{1}{\|\mathbf{f}'(t)\|}(-y'(t), x'(t))$$

Además, usando la fórmula

$$\bar{\mathbf{f}}''(s) = \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\right)$$

(obtenida en la discusión general, ver página 488), tenemos que, como $\bar{\mathbf{f}}'' = k\mathbf{N}$,

$$\frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\right) = k\mathbf{N}$$

Tomando producto punto con el vector $\mathbf{N}$ en ambos lados de esta expresión obtenemos

$$\begin{aligned}\frac{1}{\|\mathbf{f}'(t)\|^4} &= \left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t)\cdot\mathbf{N} - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\cdot\mathbf{N}\right)\\ &= k\mathbf{N}\cdot\mathbf{N} = k\|\mathbf{N}\|^2 = k\end{aligned}$$

de donde

$$\begin{aligned}k &= \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t)\cdot\mathbf{N} - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\cdot\mathbf{N}\right)\\ &= \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2(x''(t), y''(t))\cdot\frac{1}{\|\mathbf{f}'(t)\|}(-y'(t), x'(t))\right.\\ &\qquad \left. - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))(x'(t), y'(t))\cdot\frac{1}{\|\mathbf{f}'(t)\|}(-y'(t), x'(t))\right)\\ &= \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|(x'(t)y''(t) - x''(t)y'(t) - 0)\right)\\ &= \frac{x'(t)y''(t) - x''(t)y'(t)}{\|\mathbf{f}'(t)\|^3} = \frac{x'(t)y''(t) - x''(t)y'(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}}\end{aligned}$$

fórmula que coincide (excepto posiblememte por un signo) con la obtenida anteriormente considerando el camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$ como un caso particular de un camino en $\mathbb{R}^3$.

Definición. Sea $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (x(t), y(t))$ un camino regular dos veces diferenciable. Se define la curvatura (con signo) de $\mathbf{f}$ en t como

$$k(t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}}$$

■

Es interesante analizar qué tipo de información nos proporciona el signo de la curvtura $k(t)$ así definida para caminos en $\mathbb{R}^2$. Desde el punto de vista mecánico, el signo de $k(t)$ tiene que ver con la dirección hacia donde apunta el vector segunda derivada $\mathbf{f}''(t) = (x''(t), y''(t))$ (llamado "vector aceleración"). De hecho, si consideramos la curva descrita por $\mathbf{f}$, recorrida por el camino $\bar{\mathbf{f}}$ = reparametrización por longitud de arco de $\mathbf{f}$, vemos que en cada punto de la curva, el vector $\mathbf{N}(s)$, obtenido al girar el vector unitario $\mathbf{T}(s)$ en un ángulo de 90° en sentido antihorario, es un vector paralelo al vector $\mathbf{T}'(s) = \bar{\mathbf{f}}''(s)$. Este es el vector aceleración del camino $\bar{\mathbf{f}}$. Entonces, si $\mathbf{N}(s)$ y $\mathbf{T}'(s)$ coinciden en su dirección, la curvatura será positiva; en caso contrario será negativa. En general, el vector aceleración para el camino $\mathbf{f}$, i.e. el vector $\mathbf{f}''(t)$, no será paralelo al vector $\mathbf{N}(s)$; sin embargo, es posible descomponerlo como suma de sus componentes ortogonales en las direcciones de $\mathbf{T}(s)$ y $\mathbf{N}(s)$, digamos $\mathbf{f}''(t) = a(t)\mathbf{T}(s) + b(t)\mathbf{N}(s)$ donde $a = a(t)$ y $b = b(t)$ son funciones reales que dependen de t. Esta es una interesante fórmula que obtendremos en la sección 11 en la que abordaremos más detenidamente la misma mecánica del estudio de los caminos y curvas en $\mathbb{R}^2$ y $\mathbb{R}^3$. La componente $b\mathbf{N}(s)$ de $\mathbf{f}''(t)$ en la dirección de $\mathbf{N}(s)$ es a la que se aplica la condición de coincidencia con la dirección de $\mathbf{N}(s)$ (si $b > 0$) en cuyo caso la curvatura será positiva, o, en caso contrario, si $b\mathbf{N}(s)$ no apunta en la misma dirección que $\mathbf{N}(s)$, (si $b < 0$), será negativa.

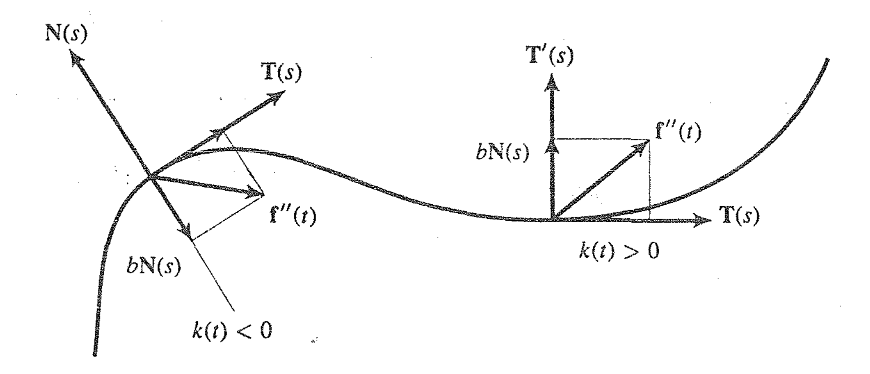


Figura 4. Curvatura con signo para un camino en $\mathbb{R}^2$.

Desde otro punto de vista, el signo de la curvatura de un camino en $\mathbb{R}^2$ tiene que ver con la concavidad hacia arriba o hacia abajo de la curva que el camino describe. Para entender esto, consideremos el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, dado por $\mathbf{f}(t) = (t, \varphi(t))$, donde $\varphi: I \subseteq \mathbb{R} \to \mathbb{R}$. Como sabemos (ver ejemplo 1 de la sección 2) la curva que describe el camino $\mathbf{f}$ es la gráfica de $y = \varphi(x)$. En este caso tenemos

$$x(t) = t, \quad x'(t) = 1, \quad x''(t) = 0$$
$$y(t) = \varphi(t), \quad y'(t) = \varphi'(t), \quad y''(t) = \varphi''(t)$$

de modo que la curvatura $k(t)$ es

$$\begin{aligned} k(t) &= \frac{x'(t)y''(t) - x''(t)y'(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}} = \frac{(1)\varphi''(t) - (0)\varphi'(t)}{(1 + (\varphi'(t))^2)^{3/2}} \\ &= \frac{\varphi''(t)}{(1 + (\varphi'(t))^2)^{3/2}} \end{aligned}$$

Es decir, que la curvatura de la función $y = \varphi(x)$ en x se calcula como

$$k(x) = \frac{\varphi''(x)}{(1 + (\varphi'(x))^2)^{3/2}}$$

De aquí se ve que el signo de $k(x)$ es el mismo que el de $\varphi''(x)$, de modo que si $k(x) > 0$ entonces la curva es cóncava hacia arriba y si $k(x) < 0$ la curva es cóncava hacia abajo. En los puntos en que $k(x) = \varphi''(x) = 0$, podría haber puntos de inflexión (donde la concavidad de la curva cambia).

Ejemplo 7. Consideremos el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (t, t^3 - 6t^2 + 9t)$. La curva que éste describe es la gráfica de la función $\varphi: \mathbb{R} \to \mathbb{R}$, $\varphi(x) = x^3 - 6x^2 + 9x$. Calculemos su curvatura. Se tiene $\varphi'(x) = 3x^2 - 12x + 9$, $\varphi''(x) = 6x - 12$. Entonces

$$k(x) = \frac{\varphi''(x)}{(1 + (\varphi'(x))^2)^{3/2}} = \frac{6x - 12}{(1 + (3x^2 - 12x + 9)^2)^{3/2}}$$

De aquí se ve que si $x < 2$, entonces $k(x) < 0$ y si $x > 2$ entonces $k(x) > 0$. en $x = 2$ la curvatura es cero. Este punto corresponde a un punto de inflexión.

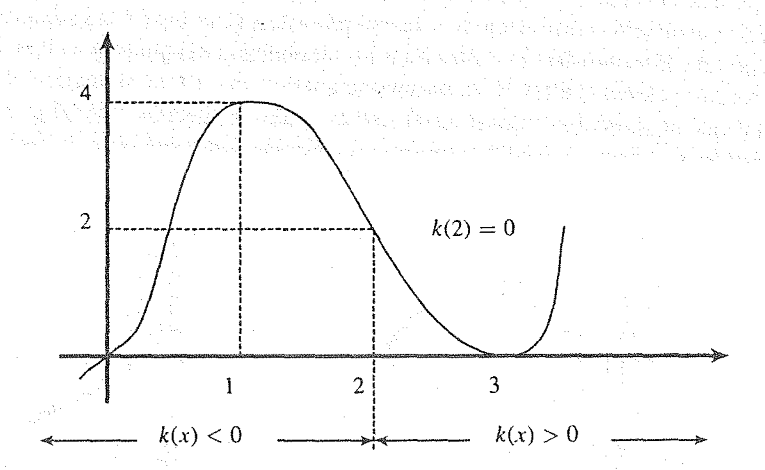


Figura 5. Gráfica de $y = x^3 - 6x^2 + 9x$ mostrando la curvatura. ■

Este tipo de conclusiones acerca de la relación entre el signo de la curvatura y la concavidad de la curva descrita por el camino, que hemos visto en el caso de que tal curva sea la gráfica de una función real de una variable, pueden ser falsas en situaciones más generales. El hecho es que no existe una relación absoluta entre *la forma* (en lo que se refiere a la concavidad) de la curva y el signo de la curvatura del camino que tiene por imagen a tal curva. Por ejemplo, si vemos que el arco de parábola $y = x^2$ entre los puntos $(-1, 1)$ y $(1, 1)$, no podemos asegurar que un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ que tenga por imagen a tal arco tendrá una curvatura positiva (pues es claro que tal arco es una curva cóncava hacia arriba). En efecto, el camino $\mathbf{f}: [-1, 1] \to \mathbb{R}^2$, $\mathbf{f}(t) = (-t, t^2)$ tiene por imagen a la curva $y = x^2$ en $-1 \leq x \leq 1$, pero en este caso

$$k(t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}} = \frac{(-1)(2) - (0)(2t)}{((-1)^2 + (2t)^2)^{3/2}} = \frac{-2}{(1 + 4t^2)^{3/2}}$$

de donde se ve que la curvatura es de hecho negativa para todo $t \in [-1, 1]$.

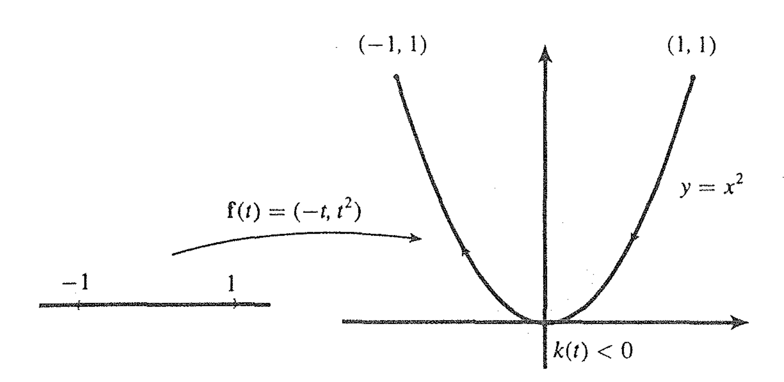


Figura 6. La curvatura del arco de parábola $y = x^2$ en $[-1, 1]$.

Obsérvese el sentido del recorrido de la curva por el camino $\mathbf{f}$. Hay una manera más consistente de abordar la relación entre el signo de la curvatura de un camino $\mathbf{f}$ y la geometría de la curva descrita por el camino. Esta consiste en rastrear en los alrededores del punto $\mathbf{p} = \mathbf{f}(t)$ el cambio en la dirección del vector velocidad $\mathbf{f}'(t)$. Si el movimiento del vector $\mathbf{f}'(t)$ (¡siguiendo el recorrido del camino $\mathbf{f}$!) es tal que su dirección cambia en el sentido de las manecillas del reloj, entonces la curvatura será negativa; si lo hace en sentido contrario de las manecillas del reloj, la curvatura será positiva

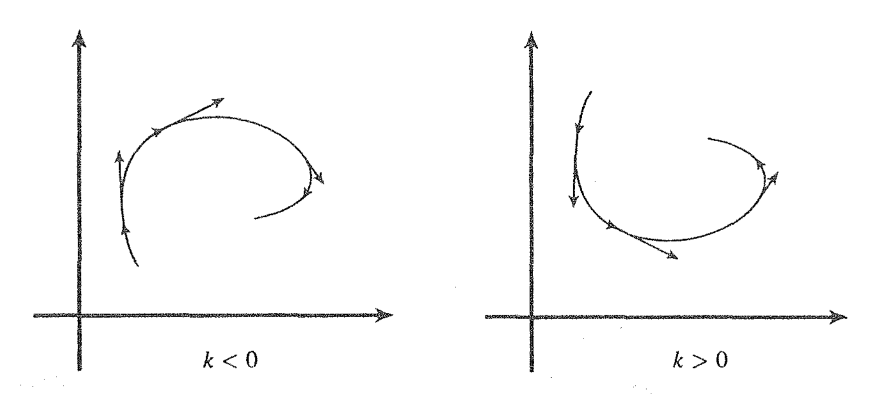


Figura 7. La curvatura con signo en relación del cambio de dirección del vector $\mathbf{f}'(t)$.

Ejemplo 8. Consideremos el camino $\mathbf{f}\colon [0, 2\pi] \to \mathbb{R}^2$ dado por $\mathbf{f}(t) = (\operatorname{sen} t, \operatorname{sen} 2t)$. La curva que éste describe es un "ocho acostado": comienza en el punto $\mathbf{f}(0) = (0, 0)$, va por el primer cuadrante (durante los primeros $\frac{\pi}{2}$ seg), pasa por $\mathbf{f}\left(\frac{\pi}{2}\right) = (1, 0)$, y regresa al origen por el cuarto cuadrante, al cual llega a los π seg ($\mathbf{f}(\pi) = (0, 0)$), pasa al segundo cuadrante y, luego de pasar por $\mathbf{f}\left(\frac{3\pi}{2}\right) = (-1, 0)$, va al tercer cuadrante, en el que está los últimos $\frac{\pi}{2}$ seg, cerrando finalmente la curva en $\mathbf{f}(2\pi) = (0, 0)$

Tenemos

$$x(t) = \operatorname{sen} t, \qquad x'(t) = \cos t, \qquad x''(t) = -\operatorname{sen} t$$
$$y(t) = \operatorname{sen} 2t, \qquad y'(t) = 2\cos 2t, \qquad y''(t) = -4\operatorname{sen} 2t$$

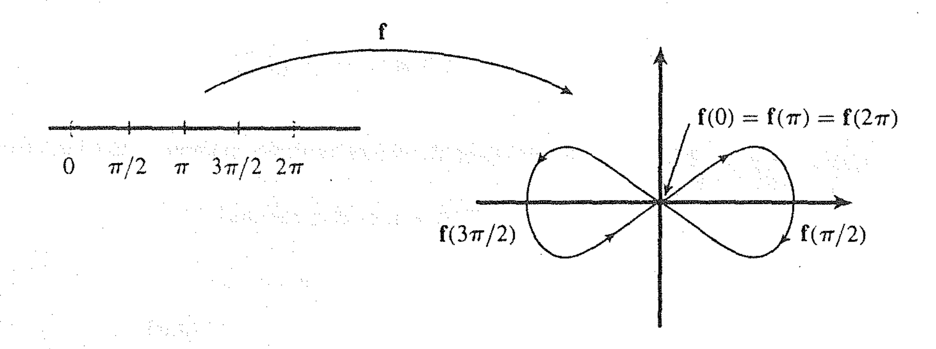


Figura 8. El camino $\mathbf{f}(t) = (\operatorname{sen} t, \operatorname{sen} 2t)$.

Entonces la curvatura es

$$k(t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}} = \frac{(\cos t)(-4\operatorname{sen} 2t) - (-\operatorname{sen} t)(2\cos 2t)}{(\cos^2 t + 4\cos^2 2t)^{3/2}}$$
$$= \frac{-2\operatorname{sen} t(\operatorname{sen}^2 t + 3\cos^2 t)}{(\cos^2 t + 4\cos^2 2t)^{3/2}}$$

Se ve que $k(t) < 0$ para $0 < t < \pi$, y $k(t) > 0$ para $\pi < t < 2\pi$. En $t = \pi$ tenemos $k(t) = 0$. Es decir, la curvatura del camino es negativa en todo el circuito cerrado derecho del ocho que describe y positiva en todo el izquierdo. Se ve claramente la correspondencia de estos signos con la variación del vector tangente al recorrer la curva.

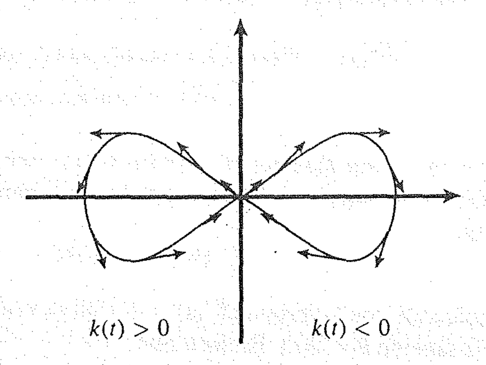


Figura 9. La curvatura del camino $\mathbf{f}(t) = (\operatorname{sen} t, \operatorname{sen} 2t)$. ■

Para entender el porqué de esta relación entre el signo de la curvatura y la dirección de variación de los vectores tangentes a la curva, recordemos que para estas curvas, imágenes de caminos $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, se definió la curvatura como el número $k(s)$ para el que $\mathbf{T}(s) = \bar{\mathbf{f}}'(s) = k(s)\mathbf{N}(s)$, donde $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^2$ es la reparametrización por longitud de arco de $\mathbf{f}$ y $\mathbf{N}(s)$ es un vector unitario que se obtiene al girar 90° el vector $\mathbf{T}(s)$ en sentido antihorario. Escribimos $\bar{\mathbf{f}}(s) = (\bar{x}(s), \bar{y}(s))$, donde

entonces $\bar{x}(s) = x(\varphi(s))$, $\bar{y}(s) = y(\varphi(s))$, con $t = \varphi(s)$ la inversa de la función

$$s = \psi(t) = \int_{t_0}^{t} \|\mathbf{f}'(u)\| du$$

Como $\bar{\mathbf{f}}'(s) = \mathbf{T}(s) = (\bar{x}'(s), \bar{y}'(s))$ es un vector unitario, podemos asociar, para cada $s \in J$ un número $\theta(s) \in \mathbb{R}$ tal

$$\mathbf{T}(s) = (\cos\theta(s), \operatorname{sen}\theta(s))$$

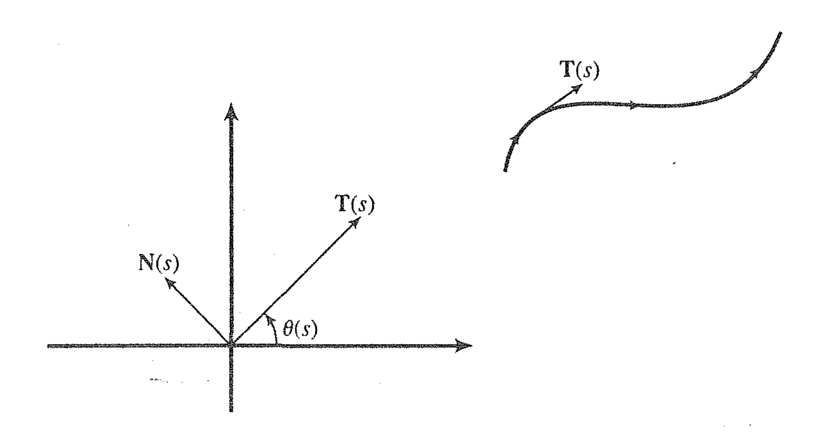


Figura 10. El vector $\mathbf{T}(s)$.

Se puede demostrar que la hipótesis hecha sobre la diferenciabilidad del camino $\mathbf{f}$ garantiza un comportamiento diferenciable para la función $\theta = \theta(s)$, de modo que

$$\begin{aligned}\bar{\mathbf{f}}''(s) = \mathbf{T}'(s) &= \big((-\operatorname{sen}\theta(s))\theta'(s), (\cos\theta(s))\theta'(s))\big) \\ &= \theta'(s)(-\operatorname{sen}\theta(s), \cos\theta(s))\end{aligned}$$

Nótese que el vector $(-\operatorname{sen}\theta(s), \cos\theta(s))$ es un vector unitario que se obtiene al girar a $\mathbf{T}(s) = (\cos\theta(s), \operatorname{sen}\theta(s))$ en un ángulo de 90° en sentido antihorario. Este es justamente el vector $\mathbf{N}(s)$. Tenemos pues que

$$\bar{\mathbf{f}}''(s) = \theta'(s)\mathbf{N}(s)$$

por lo que comparando con la fórmula $\bar{\mathbf{f}}''(s) = k(s)\mathbf{N}(s)$, vemos que la curvatura $k(s)$ es justamente la derivada de la función $\theta = \theta(s)$. Es decir que

$$k(s) = \theta'(s)$$

Ahora ya puede entender que si el vector velocidad $\mathbf{f}'(t)$ (que tiene la misma dirección que $\mathbf{T}(s)$) gira de modo que su dirección cambia en el sentido de las manecillas del reloj, entonces la función $\theta(s)$ es decreciente y por lo tanto su derivada (que es precisamente $k(s)$) es negativa. Análogamente, si el vector velocidad $\mathbf{f}'(t)$ gira cambiando su dirección en sentido antihorario, entonces la función $\theta(s)$ es creciente y por tanto su derivada (la curvatura) será positiva.

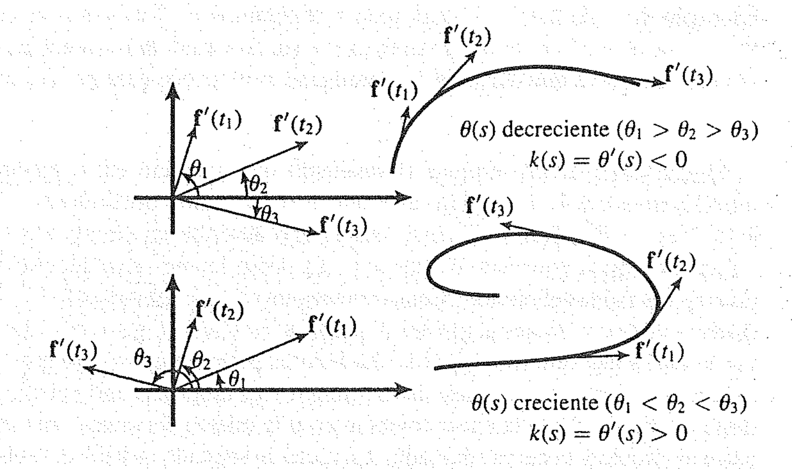


Figura 11. Una curva con $\theta(s)$ creciente y otra decreciente.

De este análisis se descubre nuevamente ante nosotros lo que es la curvatura: la fórmula $k(s) = \theta'(s)$ nos dice (como ya sabíamos) que la curvatura es la rapidez de variación de la dirección del vector velocidad de la curva. Además, este mismo análisis nos asegura el importante hecho de que las curvas en el plano están completamente determinadas por su curvatura. En otras palabras, que ante una función curvatura $k = k(s)$ existe "una única" curva en el plano que tiene curvatura $k(s)$. En efecto, suponga dada una función $k = k(s)$. Afirmamos que hay un camino $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^2$ cuya curvatura en s es precisamente $k(s)$. En efecto, según el análisis anterior tenemos que

$$\theta(s) = \int k(s)ds + \rho$$

EL vector $\mathbf{T}(s) = \bar{\mathbf{f}}'(s)$ debe ser

$$\mathbf{T}(s) = (\cos\theta(s), \operatorname{sen}\theta(s)) = (\bar{x}'(s), \bar{y}'(s))$$

de donde $\bar{\mathbf{f}}(s) = (x(s), y(s))$ debe ser tal que

$$\bar{x}(s) = \int \cos\theta(s)ds + \alpha$$

$$\bar{y}(s) = \int \operatorname{sen}\theta(s)ds + \beta$$

En resumen, el camino $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^2$ que tiene tiene en s una curvatura $k(s)$ es

$$\bar{\mathbf{f}}(s) = \left(\int \cos\theta(s)ds + \alpha, \int \operatorname{sen}\theta(s)ds + \beta\right)$$

donde

$$\theta(s) = \int k(s)ds + \rho$$

siendo α, β y ρ constantes. La aparición de éstas nos dice simplemente que, a menos que haya un *movimiento rígido* (una traslación del origen al punto (α, β) y un giro del sistema coordenado un ángulo ρ) la curva que describe el camino $\bar{\mathbf{f}}$ es única.

Ejemplo 9. Si $k(s) = 0$ para toda s, tenemos $\theta = \int 0\,ds + \rho = \text{cte} + \rho = \theta_0$, de modo que $\bar{\mathbf{f}}(s) = (\alpha, \beta) + \left(\int \cos\theta_0\,ds, \int \operatorname{sen}\theta_0\,ds\right) = (\alpha, \beta) + ((\cos\theta_0)s, (\operatorname{sen}\theta_0)s)$ es el camino cuya curvatura es cero. La curva que describe es claramente la recta que pasa por (α, β) y tiene pendiente $\tan\theta_0$. ■

Quisiéramos ahora retomar el resultado que apareció en el ejemplo 2 de esta sección para introducirnos en la discusión referente a los círculos osculadores. Ahí se vió que el camino $\mathbf{f}\colon [0, 2\pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = (r\cos t, r\operatorname{sen} t)$ que describe un círculo con centro en el origen y radio r tiene curvatura constante e igual a $\frac{1}{r}$. Es decir, la curvatura del círculo $x^2 + y^2 = r^2$ es igual al inverso del radio del círculo. Consideremos un camino arbitrario $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$ (regular, dos veces diferenciable), y digamos que en el punto $\mathbf{p} = \mathbf{f}(t)$ éste tiene curvatura $k(t) \neq 0$. Si quisiéramos ver la curva que describe $\mathbf{f}$ en los alrededores $\mathbf{p}$, como un círculo que "se confunde" con la curva, deberíamos pedir que, además de la condición de tangencia entre el círculo y la curva en el punto $\mathbf{p}$, tanto el círculo como la curva tuvieran en $\mathbf{p}$ la misma curvatura. Así tendríamos un buen contacto entre el círculo y la curva (llamado "contacto de segundo orden"; el "contacto de primer orden" es el que acontece con la recta tangente a la curva en $\mathbf{p}$ y la curva misma). El radio de este círculo debe ser (como lo sugiere el ejemplo 2 mencionado anteriormente) igual a $\frac{1}{k(s)}$. Además es claro que el centro de este círculo se debe encontrar en la dirección del vector $\bar{\mathbf{f}}''(s)$, donde $\bar{\mathbf{f}}$ es la reparametrización por longitud de arco de $\mathbf{f}$, es decir, se debe encontrar "dentro" de la curva (en torno a $\mathbf{p}$).

Establezcamos formalmente las definiciones de los conceptos anteriores.

Definición. Sea $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$ un camino regular dos veces diferenciable. Para los puntos $\mathbf{p}(t) \in \mathbb{R}^2$ de la curva que describe $\mathbf{f}$ en los cuales la curvatura $k(t)$ es no nula, se define el *radio de curvatura* de $\mathbf{f}$ en $\mathbf{p}$ denotado por $r(t)$, como $r(t) = \frac{1}{|k(t)|}$. Se llama *círculo osculador* (del latín "osculum" que significa "beso": el círculo osculador "besa" la curva en $\mathbf{p}$) de la curva en $\mathbf{p}$ al círculo que pasa por $\mathbf{p}$, tiene radio igual a $r(t)$, y cuyo centro se encuentra en la dirección del vector $\bar{\mathbf{f}}''(s)$, donde $\bar{\mathbf{f}}$ es la reparametrización por longitud de arco de $\mathbf{f}$ (es decir, si $\mathbf{c}(t) \in \mathbb{R}^2$ es el punto donde se encuentra el centro, el vector $\mathbf{c}(t) - \mathbf{f}(t)$ debe tener la misma dirección que el vector $\bar{\mathbf{f}}''(s)$).

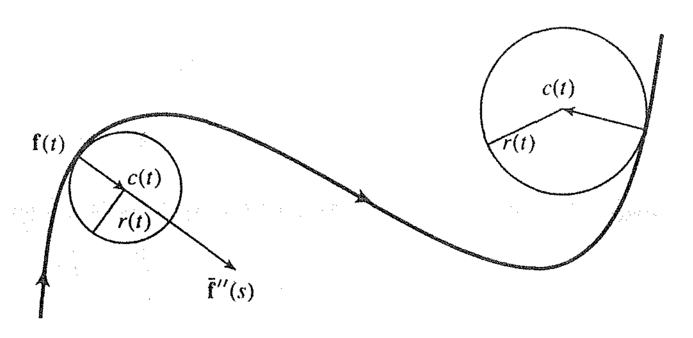


Figura 12. El círculo osculador en una curva. ■

Es claro que hay una relación inversa entre la magnitud de la curvatura $k(t)$ y el radio del círculo osculador $r(t)$, lo cual nos dice que cuanto "más plana" sea la curva (su curvatura sea menor) tanto mayor será el círculo osculador.

Obsérvese que el vector $\mathbf{c}(t) - \mathbf{f}(t)$ es tal que: **i.** tiene magnitud igual a $r(t)$ = radio de curvatura; **ii.** si $k(t) > 0$, su dirección coincide con la del vector unitario $\mathrm{N}(s) = \frac{1}{\|\mathbf{f}'(t)\|}(-y'(t), x'(t))$, pues en este caso los vectores $\bar{\mathbf{f}}''(s)$ y $\mathrm{N}(s)$ apuntan a la misma dirección; **iii.** si $k(t) < 0$, *su* dirección es opuesta a la del vector $\mathrm{N}(s)$. Estas tres condiciones se pueden escribir con la única fórmula

$$\mathbf{c}(t) - \mathbf{f}(t) = \frac{1}{k(t)}\mathrm{N}(s)$$

de donde el centro del círculo osculador (llamado *centro de curvatura)* debe estar en el punto

$$\mathbf{c}(t) = \mathbf{f}(t) + \frac{1}{k(t)}\mathrm{N}(s)$$

o mejor aún, en el punto

$$\left(x(t) - \frac{y'(t)}{k(t)\|\mathbf{f}'(t)\|}, y(t) + \frac{x'(t)}{k(t)\|\mathbf{f}'(t)\|}\right)$$

Ejemplo 10. Si $\mathbf{f}\colon [0, 2\pi] \to \mathbb{R}^2$ es el camino $\mathbf{f}(t) = (r\cos t, r\operatorname{sen} t)$, entonces $k(t) = \frac{1}{r}$, de modo que el centro del círculo osculador en $\mathbf{p} = \mathbf{f}(t)$ debe ser

$$\left(x(t) - \frac{y'(t)}{k(t)\|\mathbf{f}'(t)\|}, y(t) + \frac{x'(t)}{k(t)\|\mathbf{f}'(t)\|}\right) = \left(r\cos t - \frac{r\cos t}{\frac{1}{r}r}, r\operatorname{sen} t - \frac{r\operatorname{sen} t}{\frac{1}{r}r}\right) = (0, 0)$$

Entonces, el círculo osculador tiene en todos los puntos su centro en el origen y radio igual a $r(t) = \frac{1}{|k(t)|} = \frac{1}{\frac{1}{r}} = r$. Este es el círculo $x^2 + y^2 = r^2$ que es precisamente la imagen del camino $\mathbf{f}$. Esto nos confirma lo que era de esperarse: el círculo osculador de un círculo es él mismo. ∎

Ejemplo 11. Sea $\mathbf{f}\colon \mathbb{R}^+ \to \mathbb{R}^2$ el camino $\mathbf{f}(t) = (t, \ln t)$. La curva que este camino describe es la gráfica de la función $y = \varphi(x) = \ln x$. Calculemos su curvatura. Se tiene

$$k(x) = \frac{\varphi''(x)}{(1 + (\varphi'(x))^2)^{3/2}} = \frac{-\frac{1}{x^2}}{\left(1 + \frac{1}{x^2}\right)^{3/2}} = \frac{-x}{(1 + x^2)^{3/2}}$$

Hallemos el círculo osculador en el punto $\mathbf{p} = (1, 0)$. Tenemos que $k(1) = -\frac{1}{2\sqrt{2}}$, de modo que el radio de curvatura en $\mathbf{p}$ es $r(1) = 2\sqrt{2}$. Además

$$\|\mathbf{f}'(t)\| = \left\|\left(1, \frac{1}{t}\right)\right\| = \sqrt{1 + \frac{1}{t^2}}$$

y entonces $\|\mathbf{f}'(1)\| = \sqrt{2}$. También $x'(t) = 1$, $y'(t) = \frac{1}{t}$, de modo que $x'(1) = 1$, $y'(1) = 1$. Entonces el círculo osculador buscado debe tener su centro en

$$\left(x(t) - \frac{y'(t)}{k(t)\|\mathbf{f}'(t)\|}, y(t) + \frac{x'(t)}{k(t)\|\mathbf{f}'(t)\|}\right) = \left(1 - \frac{1}{-\frac{1}{2\sqrt{2}}\sqrt{2}}, 0 + \frac{1}{-\frac{1}{2\sqrt{2}}\sqrt{2}}\right) = (3, -2)$$

y por lo tanto tal círculo es

$$(x - 3)^2 + (y + 2)^2 = 8$$

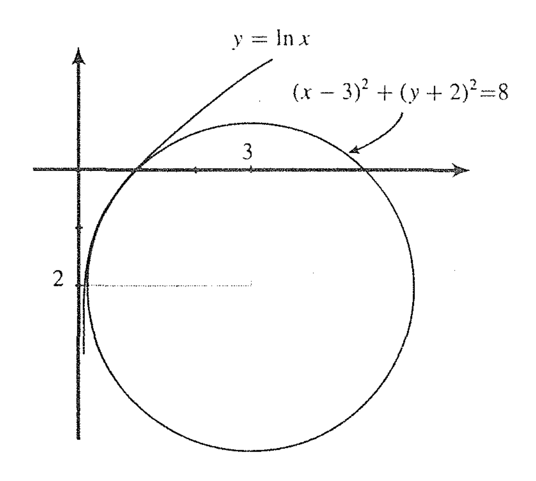


Figura 13. El camino del ejemplo 11. ■

Ejercicios (Capítulo 5, Sección 7)

En los ejercicios 1–5, determine la curvatura de la curva dada en el punto indicado.

1. $\boldsymbol{\lambda}(t) = (\cosh t, \operatorname{sen} ht, 2t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(0)$.

2. $\boldsymbol{\lambda}(t) = (e^t, e^{-t}, t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(0)$.

3. $\boldsymbol{\lambda}(t) = (t^2 + 1, t^2 - 1, t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(1)$.

4. $\boldsymbol{\lambda}(t) = (\ln t, t, t^2)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(1)$.

5. $\boldsymbol{\lambda}(t) = (t^3 - 3t, t^2 + t, 3t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(1)$.

En los ejercicios 6–10, determine la curvatura (con signo) de la curva en el plano, en un punto arbitrario de ella.

6. $\boldsymbol{\lambda}(t) = (\cosh t, \operatorname{senh} t)$.

7. $\boldsymbol{\lambda}(t) = (t^2, t^3)$.

8. $\boldsymbol{\lambda}(t) = (t^3, t^2)$.

9. $\boldsymbol{\lambda}(t) = (t^3 + t^2, t^2 - t)$.

10. $\boldsymbol{\lambda}(t) = (e^t, e^{-t})$.

11. Calcule la curvatura de la cicloide $\boldsymbol{\lambda}: \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t - \operatorname{sen} t, 1 - \cos t)$. ¿Qué sucede en puntos en los que $\cos t = 1$?

12. Calcule la curvatura de la elipse $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (a\cos t, b \operatorname{sen} t)$, en donde $0 < b < a$. ¿En qué puntos la curvatura alcanza sus valores máximo y mínimo?

13. Calcule la curvatura de la parábola $y = x^2$. ¿En qué punto de ella la curvatura alcanza su valor máximo?, ¿existe algún punto en que la curvatura sea mínima? Explique.

14. Calcule la curvatura de la parábola cúbica $y = x^3$. Estudie los extremos de la función curvatura obtenida.

15. Determine la curvatura de la parábola de n-ésimo orden $y = x^n$.

En los ejercicios 16–20 determine el círculo osculador de la curva dada en el (los) punto(s) indicado(s).

16. $y = x^2$, en el origen de coordenadas.

17. $y = \cos x$, en el punto (0, 1).

18. $y = x^3$, en el punto (1, 1) y en el punto (−1, −1).

19. $y = \dfrac{1}{1+x^2}$, en el punto (0, 1).

20. $y = e^{-x^2}$, en el punto (0, 1). (Compare con los resultados de los ejercicios 16 y 19).

21. Sea $F\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathscr{C}^2$ tal que grad $F(x, y) \neq 0\ \forall (x, y) \in U$. Si $c \in \mathbb{R}$ pertenece al rango de F, la expresión $F(x, y) = c$ define una curva en el plano cuya curvatura en un punto (x, y) de ella (demuestre que) es igual a

$$k = \frac{-\left(\dfrac{\partial F}{\partial y}\right)^2 \dfrac{\partial^2 F}{\partial x^2} + 2\dfrac{\partial F}{\partial x}\dfrac{\partial F}{\partial y}\dfrac{\partial^2 F}{\partial x \partial y} - \left(\dfrac{\partial F}{\partial x}\right)^2 \dfrac{\partial^2 F}{\partial y^2}}{\|\operatorname{grad} F\|^3}$$

22. Se llama *evoluta* de una curva $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^2$ a la curva que describen los centros de curvatura de $\mathbf{f}$.

a. Determine la evoluta de la curva $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^2)$.

b. Determine la evoluta de la curva $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, t^3)$.

(*) 5.8 Curvas paralelas

El contenido de esta sección es (con algunas modificaciones de estilo) un artículo escrito independientemente del libro (referencia [PiIII]), y motivado por una pregunta "inocente" de un estudiante de ingeniería, cuya respuesta resultó ser bastante poco inocente (este es el tipo de preguntas interesantes en matemáticas: de contenido profundo y de formulación sencilla). La pregunta fue: "¿es la curva paralela a una parábola también una parábola?". Más bien, supongamos que en cada punto de la parábola $y = x^2$ nos alejamos de ella sobre la recta normal una distancia r constante. Formamos así una nueva curva "paralela" a $y = x^2$. La figura que queda tiene al menos todo el aspecto de una nueva parábola. ¿Esta nueva curva es efectivamente una parábola? Al intentar contestar esta pregunta surgieron una gran cantidad de ideas interesantes que culminaron en lo que ahora se presenta en esta sección. La parte de la matemática que más intervino en la respuesta es la que hemos desarrollado en el presente capítulo (geometría diferencial de curvas en el plano).

En realidad, el problema que abordaremos es más ambicioso que el planteado inicialmente. Estudiaremos curvas que tienen la propiedad de mantenerse a una distancia constante de una curva dada (medida sobre la normal de esta curva en cada punto), llamadas "curvas paralelas" (a la curva dada). Exploraremos las relaciones que existen entre algunas propiedades (como la regularidad) y aspectos geométricos (como curvatura, longitud de la curva y área encerrada) de la curva dada y las correspondientes de sus curvas paralelas. Se presentan —a manera de ejemplos ilustrativos—, los estudios correspondientes a las curvas paralelas a elipses y a parábolas.

Las curvas que aparecen en esta sección se suponen infinitamente derivables (de clase $\mathscr{C}^\infty$), de modo que "curva regular" significa una curva de clase $\mathscr{C}^\infty$ cuyo vector velocidad nunca es nulo.

Consideremos la curva regular $\boldsymbol{\alpha}: I \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (x(t), y(t))$. Sea $\mathbf{n}(t)$ el vector normal unitario obtenido al girar $\boldsymbol{\alpha}'(t)$ (normalizado) un ángulo de $\pi/2$ en dirección antihoraria. Sea $\boldsymbol{\beta}: I \to \mathbb{R}^2$ la curva definida como $\boldsymbol{\beta}(t) = \boldsymbol{\alpha}(t) + r\mathbf{n}(t)$, donde r es un número real no nulo dado. Obsérvese que para cada $t \in I$ se tiene que $\|\boldsymbol{\beta}(t) - \boldsymbol{\alpha}(t)\| = |r|$, es decir $\boldsymbol{\beta}$, es una curva que se mantiene a una distancia constante —igual a $|r|$— de $\boldsymbol{\alpha}$. LLamaremos a $\boldsymbol{\beta}$ "la curva r-paralela a $\boldsymbol{\alpha}$". Se entiende que r puede ser positivo o negativo, (por razones obvias se descarta el caso $r = 0$).

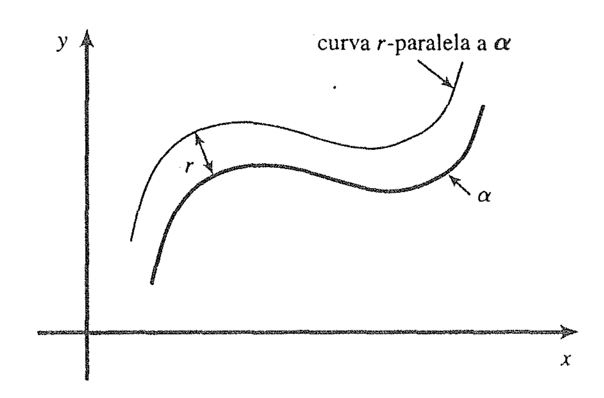


Figura 1. Curva paralela a $\boldsymbol{\alpha}$.

Comencemos por hacer explícitas las funciones coordenadas de la curva $\boldsymbol{\beta}$. Como $\boldsymbol{\alpha}'(t) = (x'(t), y'(t))$, se tiene que el vector $\mathbf{n}(t)$ es

$$\mathbf{n}(t) = \frac{1}{\|\boldsymbol{\alpha}'(t)\|}(-y'(t), x'(t))$$

de modo que

$$\boldsymbol{\beta}(t) = \boldsymbol{\alpha}(t) + r\mathbf{n}(t) = \left(x(t) - \frac{ry'(t)}{\|\boldsymbol{\alpha}'(t)\|}, y(t) + \frac{rx'(t)}{\|\boldsymbol{\alpha}'(t)\|}\right) \tag{1}$$

Obtengamos una expresión para el vector velocidad $\boldsymbol{\beta}'(t)$. Se tiene, derivando, (haciendo uso de que $(\|\boldsymbol{\rho}(t)\|)' = (\boldsymbol{\rho}(t) \cdot \boldsymbol{\rho}'(t))/\|\boldsymbol{\rho}(t)\|$ en donde $'$ denota derivación respecto de t de la norma del vector $\boldsymbol{\rho}(t)$) lo siguiente

$$\boldsymbol{\beta}'(t) = \left(\|x'(t)\| - r\frac{\|\boldsymbol{\alpha}'(t)\|y''(t) - y'(t)(\boldsymbol{\alpha}'(t) \cdot \boldsymbol{\alpha}''(t))/\|\boldsymbol{\alpha}'(t)\|}{\|\boldsymbol{\alpha}'(t)\|^2},\right.$$
$$\left. y'(t) + r\frac{\|\boldsymbol{\alpha}'(t)\|x''(t) - x'(t)(\boldsymbol{\alpha}'(t) \cdot \boldsymbol{\alpha}''(t))/\|\boldsymbol{\alpha}'(t)\|}{\|\boldsymbol{\alpha}'(t)\|^2}\right)$$

Simplificando nos queda

$$\boldsymbol{\beta}'(t) = \left(x'(t) - r\frac{x'(t)(x'(t)y''(t) - x''(t)y'(t))}{\|\boldsymbol{\alpha}'(t)\|^3}, y'(t) + r\frac{y'(t)(x''(t)y'(t) - x'(t)y''(t))}{\|\boldsymbol{\alpha}'(t)\|^3}\right)$$

o sea

$$\boldsymbol{\beta}'(t) = \left(1 - r\frac{x'(t)y''(t) - x''(t)y'(t)}{\|\boldsymbol{\alpha}'(t)\|^3}\right)(x'(t), y'(t))$$

Recordando que la curvatura con signo de la curva $\boldsymbol{\alpha}$, que denotaremos por $k_{\alpha}(t)$, está dada por

$$k_{\alpha}(t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{\|\boldsymbol{\alpha}'(t)\|^3}$$

nos queda finalmente la fórmula para $\boldsymbol{\beta}'(t)$ como

$$\boxed{\boldsymbol{\beta}'(t) = (1 - rk_{\alpha}(t))\boldsymbol{\alpha}'(t)} \qquad (2)$$

Siendo la curva $\boldsymbol{\alpha}$ regular, nos preguntamos si la curva r-paralela $\boldsymbol{\beta}$ también lo es. Con la fórmula que obtuvimos para $\boldsymbol{\beta}'(t)$ es fácil contestar esta pregunta.

Teorema 5.8.1 Sea $\boldsymbol{\alpha}: I \to \mathbb{R}^2$ una curva regular, y sea $\boldsymbol{\beta}: I \to \mathbb{R}^2$ la curva r-paralela a $\boldsymbol{\alpha}$. La curva $\boldsymbol{\beta}$ es regular si y sólo si r^{-1} no pertenece al rango de $k_{\alpha}(t)$ $(t \in I)$.

Demostración Puesto que $\|\boldsymbol{\beta}'(t)\| = |1 - rk_{\alpha}(t)|\|\boldsymbol{\alpha}'(t)\|$, y por hipótesis $\|\boldsymbol{\alpha}'(t)\| \neq 0$, se tiene que $\boldsymbol{\beta}'(t) \neq 0$ si y sólo si $1 - rk_{\alpha}(t) \neq 0\ \forall t \in I$, lo cual significa que $r^{-1} \neq k_{\alpha}(t)\ \forall t \in I$, es decir que $r^{-1} \notin$ rango de $k_{\alpha}(t)$. Q.E.D.

Por ejemplo, si consideramos la curva regular $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^2$, definida en el intervalo compacto $[a, b]$ de $\mathbb{R}$, la función curvatura $k_{\alpha}(t)$, $t \in [a, b]$ alcanzará su máximo absoluto dentro del intervalo $[a, b]$, de modo que si tomamos $r = \pm\epsilon$, $(\epsilon > 0)$, con

$$\epsilon^{-1} > \text{máx}_{t\in[a,b]} |k_{\alpha}(t)|$$

las curvas $\boldsymbol{\beta}_1(t) = \boldsymbol{\alpha}(t) + \epsilon\mathbf{n}(t)$ y $\boldsymbol{\beta}_2(t) = \boldsymbol{\alpha}(t) - \epsilon\mathbf{n}(t)$, que son curvas regulares ϵ-paralelas a $\boldsymbol{\alpha}$, (y que cada una de ellas se encuentra en "lados distintos de $\boldsymbol{\alpha}(t)$", es decir, $\boldsymbol{\beta}_1$ se encuentra en el lado marcado por la dirección del vector $\mathbf{n}(t)$, mientras que $\boldsymbol{\beta}_2$ se encuentra en el lado opuesto) constituyen la frontera de la vecindad abierta

$$V_{\epsilon}(\boldsymbol{\alpha}) = \{(x, y) \in \mathbb{R}^2 | \|(x, y) - \boldsymbol{\beta}_i(t)\| < \epsilon,\ i = 1, 2, t \in I\}$$

la cual es un ejemplo de un tipo de vecindades abiertas de curvas llamadas "vecindades tubulares".

A manera de ejemplo, consideremos la curva $\boldsymbol{\alpha}: [0, 2\pi] \to \mathbb{R}^2$, dada por $\boldsymbol{\alpha}(t) = (t, \operatorname{sen} t)$. La curva r-paralela a $\boldsymbol{\alpha}$ es $\boldsymbol{\beta}: [0, 2\pi] \to \mathbb{R}^2$

$$\boldsymbol{\beta}(t) = \left(t - (r\cos t)/(1 + \cos^2 t)^{1/2}, \operatorname{sen} t + r/(1 + \cos^2 t)^{1/2}\right)$$

La curvatura de $\boldsymbol{\alpha}$ en t es

$$k_{\alpha}(t) = -(\operatorname{sen} t)/(1 + \cos^2 t)^{3/2}$$

que alcanza su máximo en $t = 3\pi/2$, el cual vale $k_{\alpha}(3\pi/2) = 1$. Entonces, tomando $\epsilon > 0$ de modo que $\epsilon^{-1} > 1$, es decir que $\epsilon < 1$, tenemos que las curvas regulares ϵ-paralelas $\boldsymbol{\beta}(t)$ dadas en

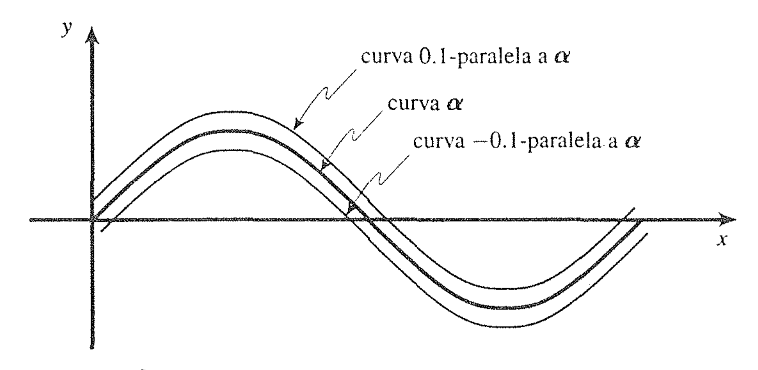


Figura 2. Curvas ± 0.1-paralelas a la gráfica de $y = \operatorname{sen} x$, $0 \leq x \leq 2\pi$.

la fórmula anterior con $r = \pm\epsilon$, constituyen la frontera de una vecindad tubular de $\boldsymbol{\alpha}$. La siguiente figura muestra el caso en que $\epsilon = 0.1$.

Veamos ahora cómo están relacionadas las curvaturas k_α, de la curva $\boldsymbol{\alpha}$, con la de su paralela $\boldsymbol{\beta}$, k_β. Tomemos $r^{-1} \notin$ rango de $k_\alpha(t)$, de modo que la curva r-paralela $\boldsymbol{\beta}$ es regular. Se tiene, derivando la expresión (2), que

$$\boldsymbol{\beta}''(t) = (1 - rk_\alpha(t))\boldsymbol{\alpha}''(t) - rk'_\alpha(t)\boldsymbol{\alpha}'(t)$$

de modo que

$$\boldsymbol{\beta}'(t) \times \boldsymbol{\beta}''(t) = \big((1 - rk_\alpha(t))\boldsymbol{\alpha}'(t)\big) \times \big((1 - rk_\alpha(t))\boldsymbol{\alpha}''(t) - rk'_\alpha(t)\boldsymbol{\alpha}'(t)\big) = (1 - rk_\alpha(t))^2\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)$$

y entonces

$$k_\beta(t) = \frac{\|\boldsymbol{\beta}'(t) \times \boldsymbol{\beta}''(t)\|}{\|\boldsymbol{\beta}'(t)\|^3} = \frac{(1 - rk_\alpha(t))^2\|\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)\|}{|1 - rk_\alpha(t)|^3\|\boldsymbol{\alpha}'(t)\|^3}$$

Por las hipótesis hechas sobre r se tiene que $1 - rk_\alpha(t) > 0 \; \forall t \in I$, o bien $1 - rk_\alpha(t) < 0 \; \forall t \in I$, y entonces la fórmula anterior queda como

$$\boxed{k_\beta(t) = \pm\frac{|k_\alpha(t)|}{1 - rk_\alpha(t)}}$$

con los signos correspondientes a cada uno de los casos mencionados. Con esta fórmula es fácil ver que si la curva $\boldsymbol{\alpha}$ tiene un vértice en t_0 (es decir que $k'_\alpha(t_0) = 0$) entonces su curva r-paralela $\boldsymbol{\beta}$ (regular) tendrá también un vértice en ese punto. En efecto, se tiene que

$$k_\beta(t) = \pm(1 - rk_\alpha(t))^{-2}k'_\beta$$

de donde se ve, de hecho, que los vértices de $\boldsymbol{\alpha}$ y de $\boldsymbol{\beta}$ coinciden.

Consideremos, a manera de ejemplo la elipse $\boldsymbol{\alpha}\colon [0, 2\pi] \to \mathbb{R}^2$ dada por $\boldsymbol{\alpha}(t) = (a\cos t, b\operatorname{sen} t)$ donde $a > b > 0$. La curvatura $k_\alpha(t)$ es

$$k_\alpha(t) = \frac{ab}{(a^2\operatorname{sen}^2 t + b^2\cos^2 t)^{3/2}}$$

Los extremos absolutos de esta función en el intervalo $[0, 2\pi]$ se alcanzan en $t = 0$, π (máximo absoluto) y en $t = \pi/2$, $3\pi/2$ (mínimo absoluto), los cuales valen $k_\alpha(0) = k_\alpha(\pi) = a/b^2$ y $k_\alpha(\pi/2) = k_\alpha(3\pi/2) = b/a^2$. Tomando entonces r tal de modo $r^{-1} \notin [b/a^2, a/b^2]$, la curva r-paralela a $\boldsymbol{\alpha}$, que según (1) es

$$\boldsymbol{\beta}(t) = \left(\cos t\left(a - \frac{rb}{(a^2 \operatorname{sen}^2 t + b^2 \cos^2 t)^{1/2}}\right), \operatorname{sen} t\left(b - \frac{ra}{(a^2 \operatorname{sen}^2 t + b^2 \cos^2 t)^{1/2}}\right)\right)$$

será una curva regular con curvatura

$$k_\beta(t) = \pm\frac{k_\alpha(t)}{1 - rk_\alpha(t)} = \pm\frac{ab}{(a^2 \operatorname{sen}^2 t + b^2 \cos^2 t)^{3/2} - rab}$$

donde el signo más corresponde al caso en el que $r < b^2/a$ (pues en este caso $r^{-1} > a/b^2 =$ máximo absoluto de $k_\alpha(t)$ en $[0, 2\pi]$, y por lo tanto se tiene que $1 - rk_\alpha(t) > 0\ \forall t \in [0, 2\pi]$), y el signo menos corresponde al caso en que $r > a^2/b$ (pues en este caso $r^{-1} < b/a^2 =$ mínimo absoluto de $k_\alpha(t)$ en $[0, 2\pi]$, y por lo tanto se tiene que $1 - rk_\alpha(t) < 0\ \forall t \in [0, 2\pi]$). La siguiente figura muestra el caso en que $a = 2$, $b = 1$, con $\boldsymbol{\beta}_1$ una curva r-paralela a $\boldsymbol{\alpha}$ siendo $r = 0.25 < 0.5 = b^2/a$, y $\boldsymbol{\beta}_2$ una curva r-paralela a $\boldsymbol{\alpha}$ con $r = 5 > 4 = a^2/b$. Obsérvese que la curva $\boldsymbol{\alpha}$ tiene curvatura positiva, al igual que las curvas $\boldsymbol{\beta}_1$ y $\boldsymbol{\beta}_2$. Sin embargo, el vector velocidad de $\boldsymbol{\beta}_1$ siempre tiene la misma dirección que el vector velocidad de $\boldsymbol{\alpha}$, en tanto que el de $\boldsymbol{\beta}_2$ siempre tiene la dirección contraria al correspondiente de $\boldsymbol{\alpha}$ (hecho que se ve fácilmente de la fórmula (2)). Nótese también que cualquier curva r-paralela a $\boldsymbol{\alpha}$ con $r < 0$, es automáticamente regular, pues siendo $k_\alpha(t) > 0\ \forall t \in [0, 2\pi]$, se tiene $1 - rk_\alpha(t) > 0\ \forall t \in [0, 2\pi]$ (figura 3).

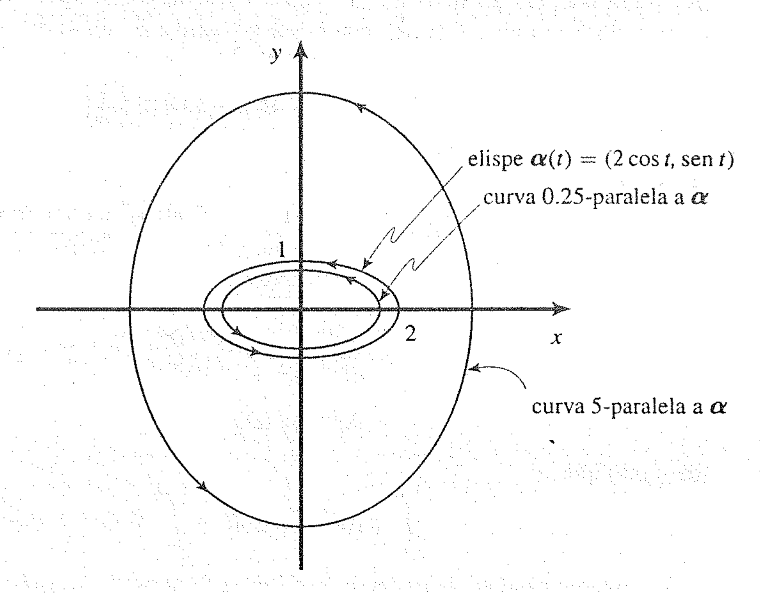


Figura 3. Curvas 0.25 y 5-paralelas a la elipse $\boldsymbol{\alpha}_t = (2\cos t, \operatorname{sen} t)$ $0 \leq t \leq 2\pi$.

Ahora estudiaremos la relación entre la longitud de la curva regular dada $\boldsymbol{\alpha}\colon [a, b] \to \mathbb{R}^2$ con la correspondiente de su curva r-paralela $\boldsymbol{\beta}\colon [a, b] \to \mathbb{R}^2$. Solamente se considerará el caso en que la

curva $\boldsymbol{\beta}$ sea regular, es decir, cuando $\boldsymbol{\beta}$ es una curva r-parelela a $\boldsymbol{\alpha}$, con r^{-1} fuera del rango de $k_{\alpha}(t)$ ($t \in [a, b]$).

Recordemos que para la curva regular $\boldsymbol{\alpha}$ se tiene bien definida su longitud (denotada como $L_{[a,b]}(\boldsymbol{\alpha})$) como

$$L_{[a,b]}(\boldsymbol{\alpha}) = \int_a^b \|\boldsymbol{\alpha}'(t)\|\, dt$$

Siendo la curva $\boldsymbol{\beta}$ también regular, su longitud se calcula como

$$L_{[a,b]}(\boldsymbol{\beta}) = \int_a^b \|\boldsymbol{\beta}'(t)\|\, dt$$

o bien, puesto que $\boldsymbol{\beta}'(t) = (1 - rk_{\alpha}(t))\boldsymbol{\alpha}'(t)$, se tiene

$$\begin{aligned} L_{[a,b]}(\boldsymbol{\beta}) &= \int_a^b |1 - rk_{\alpha}(t)| \|\boldsymbol{\alpha}'(t)\|\, dt \\ &= \int_a^b \big(\operatorname{sgn}(1 - rk_{\alpha}(t))\big)(1 - rk_{\alpha}(t))\|\boldsymbol{\alpha}'(t)\|\, dt \\ &= \operatorname{sgn}(1 - rk_{\alpha}(t))\left[\int_a^b \|\boldsymbol{\alpha}'(t)\|\, dt - r\int_a^b k_{\alpha}(t)\|\boldsymbol{\alpha}'(t)\|\, dt\right] \\ &= \operatorname{sgn}(1 - rk_{\alpha}(t))\left[L_{[a,b]} - r\int_a^b k_{\alpha}(t)\|\boldsymbol{\alpha}'(t)\|\, dt\right] \end{aligned}$$

Por otra parte, si llamamos $\theta(t)$ al ángulo formado por el vector velocidad $\boldsymbol{\alpha}'(t)$ (que por hipótesis es no nulo para toda $t \in [a, b]$) con la parte positiva del eje de las abscisas ($0 \leq \theta(t) < \pi$), se tiene

$$\theta(t) = \arctan \frac{y'(t)}{x'(t)}$$

de donde

$$\begin{aligned} \theta'(t) &= \frac{1}{1 + \left(\dfrac{y'(t)}{x'(t)}\right)^2}\left(\frac{x'(t)y''(t) - y'(t)x''(t)}{(x'(t))^2}\right) \\ &= \frac{x'(t)y''(t) - y'(t)x''(t)}{((x'(t))^2 + (y'(t))^2)^{3/2}}\left(x'(t)^2 + (y'(t))^2\right)^{1/2} \\ &= k_{\alpha}(t)\|\boldsymbol{\alpha}'(t)\| \end{aligned}$$

de modo que

$$\int_a^b k_{\alpha}(t)\|\boldsymbol{\alpha}'(t)\|\, dt = \int_a^b \theta'(t)dt = \theta(b) - \theta(a)$$

y por lo tanto finalmente podemos escribir la fórmula de $L_{[a,b]}(\boldsymbol{\beta})$ como

$$\boxed{L_{[a,b]}(\boldsymbol{\beta}) = \operatorname{sgn}(1 - rk_{\alpha}(t))\big[L_{[a,b]}(\boldsymbol{\alpha}) - r(\theta(b) - \theta(a))\big]} \tag{3}$$

Más aún, si la curva $\boldsymbol{\alpha}$ es cerrada, la diferencia $\theta(b) - \theta(a)$ es un múltiplo de 2π. En forma más precisa se tiene que $\theta(b) - \theta(a) = 2\pi I$, en I se llama índice de rotación de la curva (que de manera intuitiva lo pensamos como la cantidad de vueltas completas que da el vector tangente sobre la curva). Así pues, en este caso la fórmula (3) se ve como

$$L_{[a,b]}(\boldsymbol{\beta}) = \operatorname{sgn}(1 - rk_{\boldsymbol{\alpha}}(t))\left[L_{[a,b]}(\boldsymbol{\alpha}) - 2\pi r|I|\right]$$

de modo que si $1 - rk_{\boldsymbol{\alpha}}(t) > 0\ \forall t \in [a, b]$ (situación que se presenta si $r^{-1} > \text{máx}_{t\in[a,b]} |k_{\boldsymbol{\alpha}}(t)|$, o bien, en particular, si la curva $\boldsymbol{\alpha}$ es una curva simple convexa positivamente orientada y $r < 0$), la longitud de la curva $\boldsymbol{\beta}$, r-paralela a $\boldsymbol{\alpha}$ es

$$L_{[a,b]}(\boldsymbol{\beta}) = L_{[a,b]}(\boldsymbol{\alpha}) - 2\pi r|I|$$

en tanto que si $1 - rk_{\boldsymbol{\alpha}}(t) < 0\ \forall t \in [a, b]$ (por ejemplo, si $r^{-1} < \text{mín}_{t\in[a,b]} |k_{\boldsymbol{\alpha}}(t)|$), la curva regular $\boldsymbol{\beta}$, r-paralela a $\boldsymbol{\alpha}$ tiene por longitud

$$L_{[a,b]}(\boldsymbol{\beta}) = -L_{[a,b]}(\boldsymbol{\alpha}) + 2\pi r|I|$$

Retomando el ejemplo de la elipse $\boldsymbol{\alpha}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (2\cos t, \operatorname{sen} t)$, habíamos ya considerado curvas r-paralelas a ella con $r = 0.25$ (caso en el que $1 - rk_{\boldsymbol{\alpha}}(t) > 0\ \forall t \in [0, 2\pi]$), que llamamos $\boldsymbol{\beta}_1$, y con $r = 5$ (caso en el que $1 - rk_{\boldsymbol{\alpha}}(t) < 0\ \forall t \in [0, 2\pi]$), que llamamos $\boldsymbol{\beta}_2$ (ver figura 3). Tanto $\boldsymbol{\beta}_1$ como $\boldsymbol{\beta}_2$ son regulares y para calcular su longitud podemos usar las fórmulas recientemente obtenidas. Se tiene

$$L_{[0,2\pi]}(\boldsymbol{\beta}_1) = L_{[0,2\pi]}(\boldsymbol{\alpha}) - 2\pi r = 9.6885 - 0.5\pi = 8.1177$$
$$L_{[0,2\pi]}(\boldsymbol{\beta}_2) = -L_{[0,2\pi]}(\boldsymbol{\alpha}) + 2\pi r = -9.6885 + 10\pi = 21.7275$$

Consideremos ahora una curva regular cerrada simple $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^2$ (es decir, se tiene $\boldsymbol{\alpha}^{(j)}(a) = \boldsymbol{\alpha}^{(j)}(b)$, $j = 0, 1, 2, \ldots$, y $\boldsymbol{\alpha}(t_1) \neq \boldsymbol{\alpha}(t_2)$ para toda pareja $t_1, t_2 \in [a, b)$, $t_1 \neq t_2$). Así si tiene sentido hablar del "área encerrada por la curva $\boldsymbol{\alpha}$". De hecho, con ayuda del Teorema de Green que estudiaremos en el capítulo 7, podemos calcular dicha área (que denotaremos por $A(\boldsymbol{\alpha})$) como (con la curva $\boldsymbol{\alpha}(t) = (x(t), y(t))$ positivamente orientada)

$$A(\boldsymbol{\alpha}) = \frac{1}{2}\int_a^b [x(t)y'(t) - x'(t)y(t)]\,dt$$

(en este momento, aceptemos que esta fórmula proporciona efectivamente el área mencionada). Vamos ahora a estudiar cómo está relacionada esta área con la correspondiente de su curva r-paralela (cuando tenga sentido hablar del área de esta curva). Primero obsérvese que si la curva $\boldsymbol{\alpha}$ es cerrada entonces la curva $\boldsymbol{\beta}$ será también cerrada (hecho que se ve fácilmente por la manera como está definida $\boldsymbol{\beta}$). La propiedad de regularidad de la curva $\boldsymbol{\alpha}$, como ya se ha visto, se puede perder en la curva $\boldsymbol{\beta}$; esto depende de la existencia de raíces de la ecuación $1 - rk_{\boldsymbol{\alpha}}(t) = 0$ en $[a, b]$. Este hecho, sin embargo, no nos impide usar la fórmula del Teorema de Green para calcular el área encerrada por la curva, pues ésta sigue siendo válida para curvas regulares por secciones en $[a, b]$. Entonces la propiedad importante que debe conservar la curva $\boldsymbol{\beta}$ para poder seguir hablando del área que ésta encierra es la de ser una curva simple (sin autointersecciones). Notemos también que la orientación positiva de $\boldsymbol{\alpha}$ puede perderse en $\boldsymbol{\beta}$, aun conservando esta última la propiedad de ser simple (ver el

ejemplo de la curva $\boldsymbol{\beta}$, 1.5-paralela a la elipse $\boldsymbol{\alpha}(t) = (2\cos t, \operatorname{sen} t)$, que tiene orientación negativa). De cualquier modo, este conflicto se arregla fácilmente con una reparametrización de la curva que invierta su orientación.

Partamos entonces de una curva $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (x(t), y(t))$ regular, cerrada, simple y positivamente orientada y consideremos la curva $\boldsymbol{\beta}: [a, b] \to \mathbb{R}^2$, r-paralela a $\boldsymbol{\alpha}$, con r de modo que $\boldsymbol{\beta}$ sea simple. Supongamos también, por el momento, que $\boldsymbol{\beta}$ queda orientada positivamente. Escribiendo

$$\boldsymbol{\beta}(t) = (x_{\beta}(t), y_{\beta}(t)), \quad \boldsymbol{\beta}'(t) = (x'_{\beta}(t), y'_{\beta}(t))$$

donde entonces

$$x_{\beta}(t) = x(t) - \frac{ry'(t)}{\|\boldsymbol{\alpha}'(t)\|}, \qquad y_{\beta}(t) = y(t) + \frac{rx'(t)}{\|\boldsymbol{\alpha}'(t)\|},$$
$$x'_{\beta}(t) = (1 - rk_{\alpha}(t))x'(t) \qquad y'_{\beta}(t) = (1 - rk_{\alpha}(t))y'(t)$$

tenemos que el área encerrada por $\boldsymbol{\beta}$, que denotaremos por $A(\boldsymbol{\beta})$ es

$$\begin{aligned}
A(\boldsymbol{\beta}) &= \frac{1}{2}\int_a^b \left[x_{\beta}(t)y'_{\beta}(t) - x'_{\beta}(t)y_{\beta}(t)\right] dt \\
&= \frac{1}{2}\int_a^b \left[\left(x(t) - \frac{ry'(t)}{\|\alpha'(t)\|}\right)\big((1 - rk_{\alpha}(t))y'(t)\big) - \left(y(t) - \frac{rx'(t)}{\|\alpha'(t)\|}\right)\big((1 - rk_{\alpha}(t))x'(t)\big)\right] dt \\
&= \frac{1}{2}\int_a^b (1 - rk_{\alpha}(t))\big(x(t)y'(t) - y(t)x'(t) - r\|\alpha'(t)\|\big)\, dt \\
&= \frac{1}{2}\int_a^b \big(x(t)y'(t) - y(t)x'(t)\big)\, dt - \frac{1}{2}r\int_a^b \|\alpha'(t)\| dt \\
&\quad - \frac{1}{2}r\int_a^b \big(x(t)y'(t) - y(t)x'(t)\big)k_{\alpha}(t)\, dt + \frac{1}{2}r^2\int_a^b k_{\alpha}(t)\|\alpha'(t)\|\, dt
\end{aligned}$$

El primero, segundo y cuarto sumandos son expresiones fácilmente reconocibles. En efecto, el primero no es más que el área encerrada por la curva $\boldsymbol{\alpha}$. La integral que aparece en el segundo sumando es la longitud de la curva $\boldsymbol{\alpha}$ (que denotaremos por $L(\boldsymbol{\alpha})$), y, por último, la integral que aparece en el cuarto sumando es (como ya habíamos observado en los cálculos correspondientes a la longitud de la curva paralela $\boldsymbol{\beta}$)

$$\int_a^b k_{\alpha}(t)\|\boldsymbol{\alpha}'(t)\|\, dt = \theta(b) - \theta(a) = 2\pi$$

pues la curva $\boldsymbol{\alpha}$ es simple y positivamente orientada (su índice de rotación es $+1$). Entonces la última expresión nos queda como

$$A(\boldsymbol{\beta}) = A(\boldsymbol{\alpha}) - 0.5rL(\boldsymbol{\alpha}) + \pi r^2 - 0.5r\int_a^b \big(x(t)y'(t) - y(t)x'(t)\big)k_{\alpha}(t)dt$$

Trabajemos con la integral que aparece en esta fórmula. Llamémosla ξ. Se tiene que

$$
\begin{aligned}
\xi &= \int_a^b \big(x(t)y'(t) - y(t)x'(t)\big)k_{\alpha}(t)\,dt \\
&= \int_a^b \frac{(x(t)y'(t) - y(t)x'(t))(x'(t)y''(t) - y'(t)x''(t))}{\|\boldsymbol{\alpha}'(t)\|^3}\,dt \\
&= \int_a^b \frac{(\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}'(t))(\boldsymbol{\alpha}'(t)\cdot\boldsymbol{\alpha}''(t)) - (\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}''(t))\|\boldsymbol{\alpha}'(t)\|^2}{\|\boldsymbol{\alpha}'(t)\|^3}\,dt \\
&= \int_a^b \frac{\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}'(t)}{\|\boldsymbol{\alpha}'(t)\|^2}\int_a^b \frac{\boldsymbol{\alpha}'(t)\cdot\boldsymbol{\alpha}''(t)}{\|\boldsymbol{\alpha}'(t)\|}\,dt - \int_a^b \frac{\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}''(t)}{\|\boldsymbol{\alpha}'(t)\|}\,dt
\end{aligned}
$$

Resolviendo la primera de estas integrales (llamémosla ξ_1) por partes, tomando como u al primer factor y como dv al segundo (de modo que $v = \|\boldsymbol{\alpha}'(t)\|$). Nos queda

$$
\xi_1 = \frac{\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}'(t)}{\|\boldsymbol{\alpha}'(t)\|}\bigg]_a^b - \int_a^b \frac{\|\boldsymbol{\alpha}'(t)\|^2(\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}''(t) + \|\boldsymbol{\alpha}'(t)\|^2) - 2(\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}'(t))(\boldsymbol{\alpha}'(t)\cdot\boldsymbol{\alpha}''(t))}{\|\boldsymbol{\alpha}'(t)\|^3}\,dt
$$

Puesto que $\boldsymbol{\alpha}$ es una curva cerrada, el primer sumando de esta última expresión es igual a cero, quedándonos entonces que

$$
\xi_1 = -\int_a^b \frac{\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}''(t)}{\|\boldsymbol{\alpha}'(t)\|}\,dt - \int_a^b \|\boldsymbol{\alpha}'(t)\|\,dt + 2\int_a^b \frac{(\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}'(t))(\boldsymbol{\alpha}'(t)\cdot\boldsymbol{\alpha}''(t))}{\|\boldsymbol{\alpha}'(t)\|^3}\,dt
$$

de modo que la integral original ξ queda como

$$
\xi = \xi_1 - \int_a^b \frac{\boldsymbol{\alpha}(t)\cdot\boldsymbol{\alpha}''(t)}{\|\boldsymbol{\alpha}'(t)\|}\,dt = 2\xi - \int_a^b \|\boldsymbol{\alpha}'(t)\|\,dt
$$

de donde finalmente obtenemos

$$
\xi = \int_a^b \|\boldsymbol{\alpha}'(t)\|\,dt = L(\boldsymbol{\alpha})
$$

Sustituyendo este valor en la fórmula obtenida para $A(\boldsymbol{\beta})$ llegamos a

$$
\boxed{A(\boldsymbol{\beta}) = A(\boldsymbol{\alpha}) - rL(\boldsymbol{\alpha}) + \pi r^2}
$$

Llamamos la atención al hecho de que esta fórmula siempre se puede aplicar cuando $r < 0$, pues en este caso la curva $\boldsymbol{\beta}$ conserva las propiedades de regularidad, de ser simple y la positividad de la orientación. Cuando $r > 0$, habrá que ser cuidadosos en verificar que la curva $\boldsymbol{\beta}$ sea simple para poder aplicar la fórmula anterior (la dirección en la orientación, en caso de cambiar, se reflejará en un signo menos en el resultado del área). Tomemos como ejemplo nuevamente la elipse $\boldsymbol{\alpha}\colon [0, 2\pi] \to \mathbb{R}^2$ dada por $\boldsymbol{\alpha}(t) = (2\cos t, \operatorname{sen} t)$ y consideremos sus curvas r-paralelas $\boldsymbol{\beta}$ con $r = 0.25$, $r = 5$ (casos en que la curva $\boldsymbol{\beta}$ queda regular) y $r = 1.5$. El área encerrada por la curva 0.25-paralela a la elipse es (tomando el valor ya calculado de la longitud de $\boldsymbol{\alpha}$ como $L(\boldsymbol{\alpha}) = 9.6885$, y el conocido valor del área encerrada por la elipse de semiejes a y b, que es πab)

$$
A(\boldsymbol{\beta}) = A(\boldsymbol{\alpha}) - rL(\boldsymbol{\alpha}) + \pi r^2 = 2\pi - 0.25(9.6885) + \pi(0.25)^2 = 4.05745
$$

Para la curva 5-paralela a la elipse tenemos

$$A(\boldsymbol{\beta}) = 2\pi - 5(9.6885) + \pi(5)^2 = 36.3807$$

y para la curva 1.5-paralela a la elipse se tiene

$$A(\boldsymbol{\beta}) = 2\pi - 1.5(9.6885) + \pi(1.5)^2 = -1.18095$$

Nótese que el valor del área encerrada por esta curva 1.5-paralela a la elipse es negativo, pues, como ya habíamos comentado, la orientación que tiene esta curva es negativa. Es claro, sin embargo, que en este caso el valor absoluto de nuestro resultado es el área procurada.

En el caso particular de la elipse $\boldsymbol{\alpha}\colon [0, 2\pi] \to \mathbb{R}^2$, dada por $\boldsymbol{\alpha}(t) = (a\cos t, b\,\mathrm{sen}\, t)$, con $a > b > 0$, podemos ser más explícitos en cuanto a la utilización de la fórmula del área encerrada por sus curvas r-paralelas $\boldsymbol{\beta}\colon [0, 2\pi] \to \mathbb{R}^2$

$$\boldsymbol{\beta}(t) = \left(\cos t\left(a - \frac{rb}{(a^2\,\mathrm{sen}^2\, t + b^2\cos^2 t)^{1/2}}\right), \mathrm{sen}\, t\left(b - \frac{ra}{(a^2\,\mathrm{sen}^2\, t + b^2\cos^2 t)^{1/2}}\right)\right)$$

Estudiemos las intersecciones de esta curva con el eje x. Haciendo

$$y(t) = \mathrm{sen}\, t\left(b - \frac{ra}{(a^2\,\mathrm{sen}^2\, t + b^2\cos^2 t)^{1/2}}\right) = 0$$

Nótese que siempre hay dos raíces de esta ecuación en $[0, 2\pi)$, a saber $t = 0, \pi$. Es decir, la curva $\boldsymbol{\beta}$ corta al eje x al menos en estos dos puntos. Puede, sin embargo, haber otras raíces de la ecuación, provenientes de

$$b - \frac{ra}{(a^2\,\mathrm{sen}^2\, t + b^2\cos^2 t)^{1/2}} = 0$$

Obsérvese que si $r < 0$, estas raíces no existen (las curvas paralelas hacia el exterior de la elipse cortan solamente en $t = 0$ y $t = \pi$ al eje de las x). Tomemos entonces $r > 0$. En este caso se tiene que

$$\mathrm{sen}^2\, t = \frac{r^2a^2 - b^4}{b^2a^2 - b^4}$$

Estos valores de t existen (y son diferentes de 0 y π) si y sólo si

$$0 < \frac{r^2a^2 - b^4}{b^2a^2 - b^4} < 1$$

es decir, si $b^2/a < r < b$. En tal caso, obsérvese que existen dos raíces de la ecuación $y(t) = 0$ (además de las ya mencionadas 0 y π) entre 0 y 2π, a saber

$$t_1 = \mathrm{arcsen}\left(\frac{r^2a^2 - b^4}{b^2a^2 - b^4}\right)^{1/2}, \qquad t_2 = t_1 + \pi$$

y que además se tiene $x_{\boldsymbol{\beta}}(t_1) = x_{\boldsymbol{\beta}}(t_2)$. Esto significa que bajo la condición $b^2/a < r < b$, existe una autointersección de la curva $\boldsymbol{\beta}$, la cual ocurre sobre el eje de las x (en el punto $\boldsymbol{\beta}(t_1) = \boldsymbol{\beta}(t_2)$).

De manera análoga se ve que la curva $\boldsymbol{\beta}$ siempre tiene dos intersecciones con el eje y, en $t = \pi/2$ y en $t = 3\pi/2$, y que puede tener más, provenientes de la ecuación

$$a - \frac{rb}{(a^2 \operatorname{sen}^2 t + b^2 \cos^2 t)^{1/2}} = 0$$

de donde se ve también que de ser $r < 0$, tales nuevas intersecciones (las raíces de esta ecuación) no existen (las curvas paralelas hacia el exterior de la elipse cortan al eje de las y solamente en $t = \pi/2$ y $t = 3\pi/2$). Escribiendo la última ecuación como

$$\cos^2 t = \frac{a^4 - r^2b^2}{a^4 - a^2b^2}$$

vemos que las nuevas raíces existen si y sólo si

$$0 < \frac{a^4 - r^2b^2}{a^4 - a^2b^2} < 1$$

o sea si $a < r < a^2/b$. En tal caso, existen dos raíces, a saber

$$t_1 = \arccos\left(\frac{a^4 - r^2b^2}{a^4 - a^2b^2}\right), \qquad t_2 = t_1 + \pi$$

teniéndose además que $y_{\beta}(t_1) = y_{\beta}(t_2)$. Esto significa entonces que con la condición $a < r < a^2/b$, hay dos intersecciones en el eje y (además de las correspondientes a $t = \pi/2$ y $t = 3\pi/2$) las cuales son autointersecciones de la curva $\boldsymbol{\beta}$.

En conclusión: si $b^2/a < r < b$, la curva $\boldsymbol{\beta}$ tiene autointersecciones (en el eje x); si $a < r < a^2/b$, la curva $\boldsymbol{\beta}$ tiene autointersecciones (en el eje y). Como $a > b$, observamos que en el intervalo $b \leq r \leq a$, la curva $\boldsymbol{\beta}$ no tiene autointersecciones (nótese que en este intervalo la curva $\boldsymbol{\beta}$ no es regular pues, como habíamos ya visto, la región en la cual la curva $\boldsymbol{\beta}$ pierde su regularidad corresponde a valores de r tales que $b^2/a \leq r \leq a^2/b$). Más aún, con el valor de r en el intervalo $[b, a]$, la curva r-paralela $\boldsymbol{\beta}$, siendo cerrada, simple y regular por partes, tiene orientación negativa (un ejemplo de esta situación es la curva 1.5-paralela a la elipse $\boldsymbol{\alpha}(t) = (2\cos t, \operatorname{sen} t)$ que ya ha sido estudiada, pues en este caso $r = 1.5 \in [1, 2] = [b, a]$). A manera de resumen de las variaciones que tienen las propiedades de regularidad, simplicidad y positividad en la orientación de la curva r-paralela $\boldsymbol{\beta}$, al variar r, presentamos el esquema siguiente

<table>
<tr><td>—— 0 * ——</td><td>b^2/a * ——</td><td>b * ——</td><td>a * ——</td><td>a^2/b * —→ r</td></tr>
<tr><td rowspan="2">La curva es regular (,) y simple (,]. Tiene orientación positiva (,].</td><td colspan="3">La curva no es regular [,]</td><td rowspan="2">La curva es regular (,) y simple [,). Tiene orientación positiva [,).</td></tr>
<tr><td>La curva no es simple (autointersecciones en el eje x) (,)</td><td>La curva es simple. Tiene orientación negativa [,]</td><td>La curva no es simple (autointersecciones en el eje y) (,)</td></tr>
</table>

Las siguientes figuras ilustran cada una de estas situaciones

En consecuencia, la fórmula para calcular el área encerrada por la curva r-paralela $\boldsymbol{\beta}$ a la elipse $\boldsymbol{\alpha}$, tiene sentido cuando $r \notin (b^2/a, b) \cup (a, a^2/b)$. En todos los demás casos la curva $\boldsymbol{\beta}$ es simple,

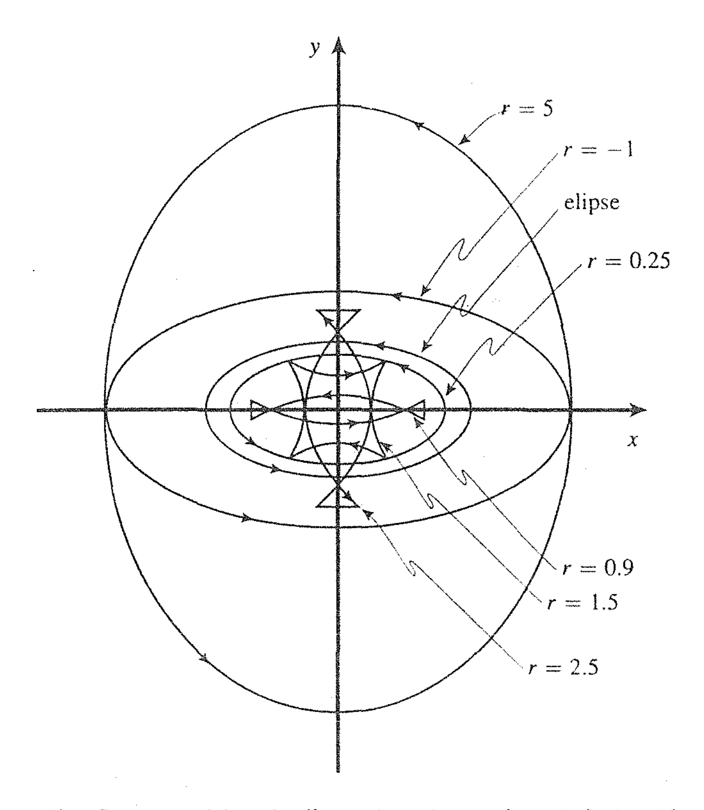


Figura 4. Curvas paralelas a la elipse $\alpha(t) = (a\cos t, b\,\mathrm{sen}\, t)$, $0 \leq t \leq 2\pi$.

propiedad fundamental para que tenga sentido hablar del área encerrada por la curva. Más aún, cuando $r \leq b^2/a$, o $r \geq a^2/b$, siendo $\boldsymbol{\beta}$ una curva positivamente orientada, el área que ésta encierra es

$$A(\boldsymbol{\beta}) = A(\boldsymbol{\alpha}) - rL(\boldsymbol{\alpha}) + \pi r^2 = \pi ab - rL(\boldsymbol{\alpha}) + \pi r^2$$

mientras que si $b \leq r \leq a$, la curva simple $\boldsymbol{\beta}$ tiene orientación negativa, y entonces el área encerrada por ella debe ser

$$A(\boldsymbol{\beta}) = -\pi ab + rL(\boldsymbol{\alpha}) - \pi r^2$$

En el teorema siguiente se recopilan los resultados obtenidos respecto a longitud y área encerradas por las curvas paralelas a una elipse.

Teorema 5.8.2 Sea $\boldsymbol{\alpha}\colon [0, 2\pi] \to \mathbb{R}^2$ la elipse

$$\boldsymbol{\alpha}(t) = (a\cos t, b\,\mathrm{sen}\, t)$$

en donde $a > b > 0$. La curva $\boldsymbol{\beta}\colon [0, 2\pi] \to \mathbb{R}^2$

$$\boldsymbol{\beta}(t) = \left(\cos t\left(a - \frac{rb}{(a^2\,\mathrm{sen}^2\, t + b^2\cos^2 t)^{1/2}}\right), \mathrm{sen}\, t\left(b - \frac{ra}{(a^2\,\mathrm{sen}^2\, t + b^2\cos^2 t)^{1/2}}\right)\right)$$

es una curva r-paralela a $\boldsymbol{\alpha}$. Sean $L(\boldsymbol{\alpha})$ y $L(\boldsymbol{\beta})$ la longitud de las curvas $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$, y $A(\boldsymbol{\alpha})$ y $A(\boldsymbol{\beta})$ el área que éstas encierran, respectivamente.

a. Si $r \leq b^2/a$, la longitud de la curva $\boldsymbol{\beta}$ se calcula como

$$L(\boldsymbol{\beta}) = L(\boldsymbol{\alpha}) - 2\pi r$$

b. Si $r \geq a^2/b$, la longitud de la curva $\boldsymbol{\beta}$ se calcula como

$$L(\boldsymbol{\beta}) = -L(\boldsymbol{\alpha}) + 2\pi r$$

c. Si $r \in (-\infty, b^2/a] \cup [a^2/b, +\infty)$, el área encerrada por la

$$A(\boldsymbol{\beta}) = \pi ab - rL(\boldsymbol{\alpha}) + \pi r^2$$

d. Si $r \in [b, a]$, el área encerrada por la curva $\boldsymbol{\beta}$ es

$$A(\boldsymbol{\beta}) = -\pi ab + rL(\boldsymbol{\alpha}) - \pi r^2$$ ■

Para finalizar esta sección, estudiaremos las curvas paralelas a una parábola. Consideremos la parábola $\boldsymbol{\alpha}: I \longrightarrow \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, pt^2)$, donde I es un intervalo de $\mathbb{R}$ simétrico respecto del origen, y p es un real positivo dado. Según la fórmula (1), la curva $\boldsymbol{\beta}$, r-paralela a $\boldsymbol{\alpha}$ es

$$\boldsymbol{\beta}(t) = \left(t - \frac{2rpt}{(1+4p^2t^2)^{1/2}}, t^2 + \frac{r}{(1+4p^2t^2)^{1/2}}\right)$$

Ciertamente ésta es una curva simétrica respecto del eje de ordenadas. Una pregunta natural sobre ella es: ¿esta curva —paralela a la parábola $y = px^2$—, también es una parábola? (Como se dijo al inicio, esta pregunta fue la que motivó en principio el presente estudio). No es difícil imaginar que la respuesta a la pregunta será negativa, pues, como hemos visto ya en varias ocasiones, la curva $\boldsymbol{\beta}$ puede llegar a perder su regularidad (propiedad que debe poseer cualquier curva parabólica). En efecto, la curvatura de $\boldsymbol{\alpha}$ en t viene dada por

$$k_{\alpha}(t) = \frac{2p}{(1+4p^2t^2)^{3/2}}$$

El rango de esta función es (pensando que $\boldsymbol{\alpha}$ está definida en todo $\mathbb{R}$) el intervalo $(0, 2p]$, de modo que, según el teorema 1, tomando $r^{-1} \in (0, 2p]$, la curva $\boldsymbol{\beta}$, r-paralela a $\boldsymbol{\alpha}$, no es regular, y no se puede esperar que ésta sea una parábola. Sin embargo, la inquietud puede persistir al imaginar una curva r-paralela a $\boldsymbol{\alpha}$ con $r < 0$ muy pequeño en valor absoluto (con $\boldsymbol{\beta}$ "muy cerca de la parte exterior de $\boldsymbol{\alpha}$"). En este caso, sin embargo, la respuesta sigue siendo negativa. En efecto, si la curva $\boldsymbol{\beta}$ fuera una parábola, ésta tendría que ser del tipo $y = \mu x^2 + \nu$, para ciertas constantes μ y ν. (En efecto, la curva $\boldsymbol{\beta}$ es simétrica respecto del eje de ordenadas y posee un extremo local en $\boldsymbol{\beta}(0) = (0, r)$, el cual corresponde al punto que está a r unidades del vértice de la parábola —que es, además, el único vértice de la curva, i.e. el punto en que $k'_{\alpha}(t) = 0$. Así, la única curva parabólica con estas características es $y = \mu x^2 + \nu$). Para probar nuestra afirmación ("la curva $\boldsymbol{\beta}$ no es una parábola") usaremos la siguiente caracterización de este tipo de curvas.

Lema La (traza de la) curva regular $\boldsymbol{\alpha}: \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (x(t), y(t))$, simétrica respecto del eje de ordenadas, es una parábola del tipo $y = \mu x^2 + \nu$, si y sólo si

$$y'(t)(x'(t))^2 - x(t)(x'(t)y''(t) - x''(t)y'(t)) = 0$$

Demostración. Primeramente veremos que la expresión anterior es invariante por reparametrizaciones de $\boldsymbol{\alpha}$. Sea $\tilde{\boldsymbol{\alpha}}: \mathbb{R} \to \mathbb{R}^2$ una reparametrización de $\boldsymbol{\alpha}$. Entonces $\tilde{\boldsymbol{\alpha}}(t) = \boldsymbol{\alpha}(\phi(t))$ para alguna función $\phi: \mathbb{R} \to \mathbb{R}$ diferenciable con derivada nunca nula en $\mathbb{R}$. Se tiene que

$$\tilde{\boldsymbol{\alpha}}'(t) = \phi'(t)\boldsymbol{\alpha}'(\phi(t)), \qquad \tilde{\boldsymbol{\alpha}}''(t) = \phi''(t)\boldsymbol{\alpha}'(\phi(t)) + (\phi'(t))^2\boldsymbol{\alpha}''(\phi(t))$$

de modo que, escribiendo $\tilde{\boldsymbol{\alpha}}(t) = (\tilde{x}(t), \tilde{y}(t))$, se tiene

$$\begin{aligned}\tilde{y}'(t)\big(\tilde{x}'(t)\big)^2 - \tilde{x}(t)\big(\tilde{x}'(t)\tilde{y}''(t) - \tilde{x}''(t)\tilde{y}'(t)\big) &= \phi'(t)y'(\phi(t))[\phi'(t)x'(\phi(t))]^2 \\ &- x(\phi(t))\big\{\phi'(t)x'(\phi(t))\big[\phi''(t)y'(\phi(t)) + (\phi'(t))^2y''(\phi(t))\big] \\ &- \phi'(t)y'(\phi(t))\big[\phi''(t)x'(\phi(t)) + (\phi'(t))^2x''(\phi(t))\big]\big\} \\ &(\phi'(t))^3\big[y'(t)(x'(t))^2 - x(t)(x'(t)y''(t) - x''(t)y'(t))\big] = 0\end{aligned}$$

Supongamos entonces que la (traza de la) curva $\boldsymbol{\alpha}$ es de la forma mencionada en el lema. Es decir, supongamos que $y(t) = \mu(x(t))^2 + \nu$, $\forall t \in \mathbb{R}$. Tomemos $x(t) = t$, y entonces $y(t) = \mu t^2 + \nu$. Se tiene

$$y'(t)(x'(t))^2 - x(t)(x'(t)y''(t) - x''(t)y'(t)) = (2\mu t) - t(2\mu) = 0$$

En forma recíproca, suponga que se cumple la relación dada. Poniendo $x(t) = t$, nos queda que $y(t)$ satisface

$$y'(t) - t(y''(t)) = 0$$

de donde, integrando nos queda

$$y(t) = \mu t^2 + \nu$$

como queríamos. Q.E.D.

Ahora es fácil ver que nuestra curva $\boldsymbol{\beta}$, r-paralela a la parábola $\boldsymbol{\alpha}$, no es una parábola del tipo $y = \mu x^2 + \nu$. De hecho se tiene (escribiendo las funciones coordenadas de $\boldsymbol{\beta}$ como $x_\beta(t)$ y $y_\beta(t)$)

$$y'_\beta(t)(x'_\beta(t))^2 - x_\beta(t)(x'_\beta(t)y''_\beta(t) - x''_\beta(t)y'_\beta(t)) = 8rp^3t^3k_\alpha(t)(1 - rk_\alpha(t))^2 \neq 0$$

Si consideramos la parábola $y = px^2$ en todo su dominio, es decir, la (traza de la) curva $\boldsymbol{\alpha}: \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, pt^2)$, como dijimos anteriormente, la curva $\boldsymbol{\beta}$, r-paralela a $\boldsymbol{\alpha}$, será regular si y sólo si $r^{-1} \notin (0, 2p] =$ rango de $k_\alpha(t)$. Es decir, $\boldsymbol{\beta}$ es regular si y sólo si $r < 1/2p$. Esto significa que, globalmente, sólo se pueden tener curvas regulares paralelas a la parábola $y = px^2$ por afuera de ella ($r < 0$), o bien, por adentro hasta $r = 1/2p$ (no incluso, o sea $0 < r < 1/2p$). Para $r \geq 1/2p$, la curva $\boldsymbol{\beta}$ no es regular, teniéndose de hecho que si $r > 1/2p$, esta curva presentará una autointersección sobre el eje de las ordenadas. En efecto, haciendo $x_\beta(t) = 0$, es decir

$$t - \frac{2rpt}{(1 + 4p^2t^2)^{1/2}} = 0$$

vemos que una raíz de esta ecuación es siempre $t = 0$ (la curva $\boldsymbol{\beta}$ siempre pasa por el eje de ordenadas en $\boldsymbol{\beta}(0) = (0, r)$), teniéndose además la posibilidad de que

$$1 - \frac{2rpt}{(1 + 4p^2t^2)^{1/2}} = 0$$

(Nótese que esta ecuación no tiene solución si $r < 0$, es decir, las curvas r-paralelas hacia el exterior de la parábola solamente cruzan el eje de ordenadas en el punto $(0, r)$), de donde

$$\tilde{t}_{1,2} = \pm(1/2p)(4r^2p^2 - 1)^{1/2}$$

(nótese también que si $r < 1/2p$ estas raíces no existen). Para $r > 1/2p$, éstas son dos raíces reales distintas en las que (puesto que $y_{\beta}(t)$ es una función par) $y_{\beta}(\tilde{t}_1) = y_{\beta}(\tilde{t}_2)$. Estas raíces corresponden entonces a una autointersección de la curva $\boldsymbol{\beta}$ sobre el eje de ordenadas. Para $r = 1/2p$ (las raíces $\tilde{t}_{1,2}$ son iguales —a cero), la curva $\boldsymbol{\beta}$, aunque no regular (de hecho $\boldsymbol{\beta}'(0) = 0$), es simple.

Veamos un poco más de cerca qué ocurre cuando $r \geq 1/2p$ (es decir, cuando la curva $\boldsymbol{\beta}$ no es regular). En este caso la ecuación

$$1 - rk_{\alpha}(t) = 1 - \frac{2pr}{(1 + 4p^2t^2)^{3/2}} = 0$$

posee dos raíces en

$$t^*_{1,2} = \pm(1/2p)\left((2pr)^{2/3} - 1\right)^{1/2}$$

Nuevamente observamos que si $r = 1/2p$, estas raíces son iguales, y la curva $\boldsymbol{\beta}$, aunque no regular (pierde su regularidad únicamente en $t_1^* = t_2^*$), es una curva simple (en este caso todas las raíces $\tilde{t}_1$, $\tilde{t}_2$, t_1^*, t_2^* son iguales a cero). Cuando $r > 1/2p$, las raíces t_1^* y t_2^* son distintas, y se tiene además que $\tilde{t}_1 < t_1^* < t_2^* < \tilde{t}_2$. Por la simetría de la curva $\boldsymbol{\beta}$ respecto del eje de ordenadas, podemos concluír que la parte de la curva $\boldsymbol{\beta}$ entre $\tilde{t}_1$ y $\tilde{t}_2$, corresponde a un "loop" (simétrico respecto del eje de ordenadas) dentro del cual la curva pierde su regularidad en los puntos $\boldsymbol{\beta}(t_1^*)$ y $\boldsymbol{\beta}(t_2^*)$.

La figura 5 (tomando $p = 1$) ilustra los casos en los que $r = -1$ (curva regular paralela hacia el exterior de la parábola $y = x^2$), $r = 0.25 < 0.5 = 1/2p$ (curva regular paralela hacia el interior de la parábola) y $r = 1.8 > 1/2p$ (curva no regular paralela hacia el interior de la parábola).

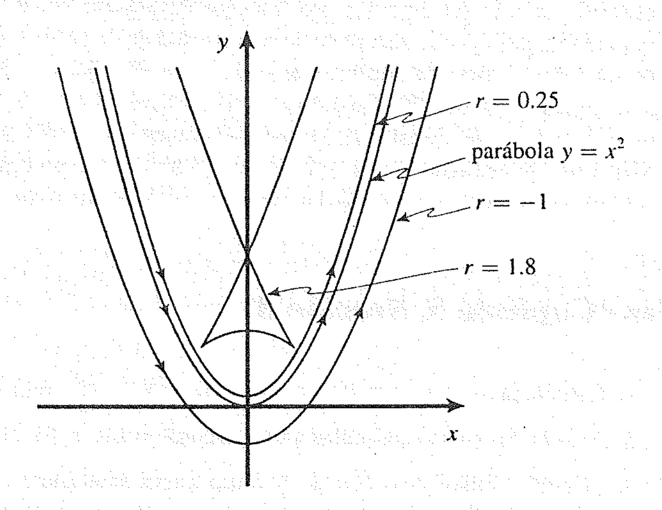


Figura 5. La parábola $y = x^2$ y sus curvas r-paralelas con $r = -1, 0.25, 1.8$.

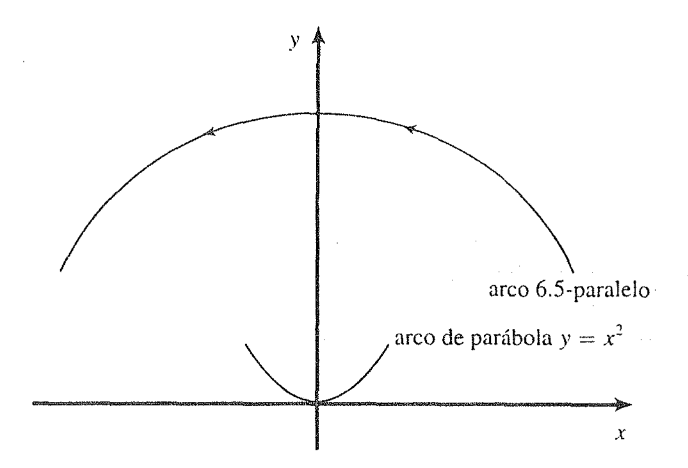


Figura 6. El arco de parábola $y = x^2$, $-1 \leq x \leq 1$ y su curva 6.5-paralela.

Si consideramos solamente un arco de parábola (simétrico respecto del eje de ordenadas), digamos $\boldsymbol{\alpha}: [-a, a] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, pt^2)$, es posible hacer algunas observaciones adicionales a las anteriormente hechas en el caso general. En este caso la función curvatura $k_{\alpha}: [-a, a] \to \mathbb{R}$

$$k_{\alpha}(t) = \frac{2p}{(1 + 4p^2t^2)^{3/2}}$$

alcanza su máximo absoluto en $t = 0$, que vale $k_{\alpha}(0) = 2p$, y alcanza también su mínimo absoluto en $t = \pm a$, que vale $k_{\alpha}(\pm a) = 2p/(1 + 4p^2a^2)^{3/2}$. LLamemos q a este valor. Se tiene entonces que la curva $\boldsymbol{\beta}$ será regular si y sólo si $r^{-1} \notin [q, 2p]$, es decir, si $r < 1/2p$ (como en el caso general) o si $r > 1/q = (1 + 4p^2a^2)^{3/2}/2p$. Este valor de q es positivo, lo indica la curva r-paralela $\boldsymbol{\beta}$ que está en la parte "interna" de la parábola. Es posible entonces tener una "imagen paralela al arco de parábola $y = px^2$, que sea regular". Observe sin embargo que el valor $1/q$ es en general muy grande (del orden de $4a^3$), lo que significa que "hay que despegarse mucho —hacia adentro, paralelamente— de la parábola para poder recuperar la regularidad de la curva r-paralela $\boldsymbol{\beta}$". Consideremos a manera de ejemplo el arco de parábola $\boldsymbol{\alpha}: [-1, 1] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, t^2)$. En este caso $1/2p = 1/2$ y $q = 2/(5)^{3/2} = 0.1789$. Cuando $r < 1/2$, la curva $\boldsymbol{\beta}: [-1, 1] \to \mathbb{R}^2$, r-paralela a $\boldsymbol{\alpha}$, es regular, estando ésta en el exterior de la parábola cuando $r < 0$ y en su interior cuando $0 < r < 1/2$. Mientras r se encuentre entre $1/2$ y $1/q = 5.59016$, la curva $\boldsymbol{\beta}$ (en el interior de la parábola) no será regular y ésta vuelve a ser regular para $r > 5.59016$. La figura 6 ilustra esta situación.

Ejercicios (Capítulo 5, Sección 8)

1. Estudie las curvas paralelas a una recta $\boldsymbol{\lambda}: \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, mt + b)$.

2. Estudie las curvas paralelas a una circunferencia $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (a \cos t, a \operatorname{sen} t)$.

(***) **3.** ¿Cierto o falso? Sea $F(x, y) = 0$ una curva en el plano. Considere sus curvas ϵ-paralelas con ϵ pequeño. Estas pueden escribirse como $G(x, y) = 0$ si y sólo si la curva $F(x, y) = 0$ es una recta o una circunferencia.

5.9 Plano osculador, normal y rectificante

Sea $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^3$ un camino regular dos veces diferenciable reparametrizado por longitud de arco. Como sabemos el vector $\mathbf{T}(s) = \mathbf{f}'(s)$ es un vector unitario tangente a la curva que describe $\mathbf{f}$ en el punto $\mathbf{p} = \mathbf{f}(s)$. Como también habíamos observado, del hecho de que el vector $\mathbf{T}(s)$ siempre tiene norma 1, se puede deducir que el vector $\mathbf{f}''(s)$ es ortogonal a $\mathbf{f}'(s)$ (en efecto, como $\|\mathbf{T}(s)\| = 1$, obtenemos, derivando, que $0 = \frac{\mathbf{T}(s)\cdot\mathbf{T}'(s)}{\|\mathbf{T}(s)\|} = \mathbf{T}(s)\cdot\mathbf{T}'(s)$, de donde se deduce la afirmación hecha). Supongamos que la curvatura de $\mathbf{f}$ en s es no nula. Podemos entonces asociar un vector normal unitario $\mathbf{N}(s)$ en la dirección de $\mathbf{f}''(s)$. De hecho se tiene $\mathbf{f}''(s) = k(s)\mathbf{N}(s)$, donde $k(s) \neq 0$ es la curvatura de $\mathbf{f}$ en s. El vector $\mathbf{N}(s)$ es llamado *vector normal principal* de $\mathbf{f}$ en s.

Los vectores $\mathbf{T}(s)$ y $\mathbf{N}(s)$ determinan (en $\mathbf{p}$) un plano, llamado *plano osculador* de $\mathbf{f}$ en s. Un vector ortogonal al plano osculador de $\mathbf{f}$ en s es el vector $\mathbf{B}(s) = \mathbf{T}(s) \times \mathbf{N}(s)$, llamado *vector binormal* de $\mathbf{f}$ en s. Obsérvese que $\mathbf{B}(s)$ es (por definición) un vector *unitario*, i.e. $\|\mathbf{B}(s)\| = 1$ (si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^3$ se tiene $\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u}\cdot\mathbf{v})^2$. Siendo $\mathbf{u}$ y $\mathbf{v}$ vectores unitarios y ortogonales se tiene $\|\mathbf{u}\times\mathbf{v}\| = (1)^2(1)^2 - 0 = 1$). Entonces el punto $\mathbf{q} = (x, y, z) \in \mathbb{R}^3$ está en el plano osculador de $\mathbf{f}$ si y sólo si el vector $\mathbf{q} - \mathbf{f}(s)$ es ortogonal a $\mathbf{B}(s)$. Es decir, la ecuación del plano osculador de $\mathbf{f}$ en s es

$$(q - \mathbf{f}(s)) \cdot \mathbf{B}(s) = 0$$

El plano atravesado por la curva descrita por $\mathbf{f}$ en forma perpendicular se le llama *plano normal* de $\mathbf{f}$ en s. Es decir, el plano normal de $\mathbf{f}$ en s es el plano que pasa por $\mathbf{f}(s)$ y tiene por vector normal a $\mathbf{T}(s)$. Se llama *plano rectificante* de $\mathbf{f}$ en s al plano que pasa por $\mathbf{p}$ y es ortogonal al plano osculador y al plano normal de $\mathbf{f}$ en s. O sea, el plano rectificante de $\mathbf{f}$ en s es el plano que pasa por $\mathbf{p}$ y tiene por vector normal a $\mathbf{N}(s)$ (ó $\mathbf{f}''(s)$).

A las rectas en $\mathbb{R}^3$ que pasan por $\mathbf{p} = \mathbf{f}(s)$ y tienen por vectores paralelos a $\mathbf{T}(s)$, $\mathbf{N}(s)$ y $\mathbf{B}(s)$ se les llama, respectivamente, recta tangente, recta normal y recta binormal de $\mathbf{f}$ en s.

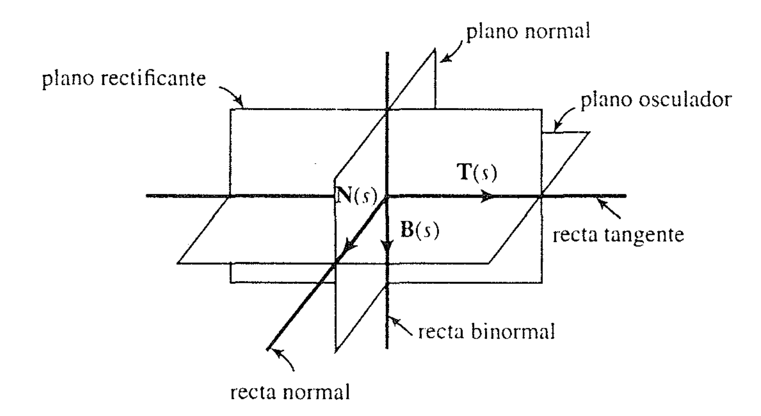


Figura 1. El plano osculador, normal y rectificante, y las rectas tangente, normal y binormal.

Es claro que las ecuaciones del plano normal y del plano rectificante de $\mathbf{f}$ en s son

$$(q - \mathbf{f}(s)) \cdot \mathbf{T}(s) = 0$$

y

$$(q - \mathbf{f}(s)) \cdot \mathbf{N}(s) = 0$$

donde $\mathbf{q} = (x, y, z)$, respectivamente. De modo semejante, las ecuaciones de la recta tangente, normal y binormal de $\mathbf{f}$ en s son

$$q = \mathbf{f}(s) + t \cdot \mathbf{T}(s)$$
$$q = \mathbf{f}(s) + t \cdot \mathbf{N}(s)$$
$$q = \mathbf{f}(s) + t \cdot \mathbf{B}(s)$$

donde $\mathbf{q} = (x, y, z)$, $t \in \mathbb{R}$, respectivamente.

Ejemplo 1. Consideremos el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por

$$\mathbf{f}(s) = \left(\cos \frac{s}{\sqrt{2}}, \operatorname{sen} \frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}\right)$$

(ver ejemplo 4 de la sección 6). Este es un camino regular, dos veces diferenciable parametrizado por longitud de arco y que describe una hélice circular en $\mathbb{R}^3$. Obtengamos la ecuaciones de los planos osculador, normal y rectificante, y las ecuaciones de las rectas tangente, normal y binormal de esta curva en el punto $\mathbf{p} = \mathbf{f}(\sqrt{2}\pi) = (0, 1, \pi)$. Tenemos

$$\mathbf{T}(s) = \mathbf{f}'(s) = \left(-\frac{1}{\sqrt{2}} \operatorname{sen} \frac{s}{\sqrt{2}}, \frac{1}{\sqrt{2}} \cos \frac{s}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

$$\mathbf{f}''(s) = \left(-\frac{1}{2} \cos \frac{s}{\sqrt{2}}, -\frac{1}{2} \operatorname{sen} \frac{s}{\sqrt{2}}, 0\right)$$

La curvatura de $\mathbf{f}$ en s es

$$k(s) = \|\mathbf{f}''(s)\| = \frac{1}{2} \neq 0$$

Entonces el vector normal unitario $\mathbf{N}(s)$ es

$$\mathbf{N}(s) = \frac{1}{k(s)} \mathbf{f}''(s) = \frac{1}{\frac{1}{2}} \left(-\frac{1}{2} \cos \frac{s}{\sqrt{2}}, -\frac{1}{2} \operatorname{sen} \frac{s}{\sqrt{2}}, 0\right)$$
$$= \left(-\cos \frac{s}{\sqrt{2}}, -\operatorname{sen} \frac{s}{\sqrt{2}}, 0\right)$$

El vector binormal $\mathbf{B}(s)$ es

$$\mathbf{B}(s) = \mathbf{T}(s) \times \mathbf{N}(s) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\frac{1}{\sqrt{2}} \operatorname{sen} \frac{s}{\sqrt{2}} & \frac{1}{\sqrt{2}} \cos \frac{s}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\cos \frac{s}{\sqrt{2}} & -\operatorname{sen} \frac{s}{\sqrt{2}} & 0 \end{bmatrix}$$
$$= \left(\frac{1}{\sqrt{2}} \operatorname{sen} \frac{s}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \cos \frac{s}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

En el punto **p** tenemos que

$$\mathbf{T}(\sqrt{2}\pi) = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \qquad \mathbf{N}(\sqrt{2}\pi) = (1, 0, 0)$$

$$\mathbf{B}(\sqrt{2}\pi) = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

(observe que estos son tres vectores unitarios ortogonales dos a dos).

El plano osculador procurado pasa por $\mathbf{p} = (0, 1, \pi)$ y tiene a $\mathbf{B}(\sqrt{2}\pi) = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ como vector normal. Su ecuación es

$$0(x - 0) + \frac{1}{\sqrt{2}}(y - 1) + \frac{1}{\sqrt{2}}(z - \pi) = 0$$

o sea

$$y + z = \pi + 1$$

El plano normal pasa por $\mathbf{p} = (0, 1, \pi)$ y tiene a $\mathbf{T}(\sqrt{2}\pi) = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ por vector normal. Su ecuación es

$$0(x - 0) - \frac{1}{\sqrt{2}}(y - 1) + \frac{1}{\sqrt{2}}(z - \pi) = 0$$

o sea

$$-y + z = \pi - 1$$

El plano rectificante pasa por $\mathbf{p} = (0, 1, \pi)$ y tiene a $\mathbf{N}(\sqrt{2}\pi) = (1, 0, 0)$ por vector normal. Su ecuación es

$$1(x - 0) + 0(y - 1) + 0(z - \pi) = 0$$

o sea

$$x = 0$$

La recta tangente a la hélice en **p** es

$$(x, y, z) = (0, 1, \pi) + t\left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

la recta normal es

$$(x, y, z) = (0, 1, \pi) + t(1, 0, 0)$$

y la recta binormal es

$$(x, y, z) = (0, 1, \pi) + t\left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$ ■

Supongamos ahora que el camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ regular dos veces diferenciable no está parametrizado por longitud de arco. Veamos cómo determinar en este caso las ecuaciones de los planos osculador, normal y rectificante a **f** en un punto $\mathbf{p} = \mathbf{f}(t)$, donde $k(t) \neq 0$, así como las ecuaciones de las rectas tangente, normal y binormal.

Sea $\bar{\mathbf{f}}: J \subseteq \mathbb{R} \to \mathbb{R}^3$, $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$, la reparametrización por longitud de arco de **f**. El vector velocidad de $\mathbf{f}'(t)$ está en la misma dirección del vector $\mathbf{T}(s) = \bar{\mathbf{f}}'(s)$ en el punto $\mathbf{p} = \mathbf{f}(t) = \bar{\mathbf{f}}(s)$,

donde $t = \varphi(s)$, así que con él podemos obtener la ecuación de la recta tangente y el plano normal de $\mathbf{f}$ en t. Se tienen las siguientes relaciones

$$\bar{\mathbf{f}}'(s) = \frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t)$$
$$\bar{\mathbf{f}}''(s) = \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\right)$$

(ver página 488). De estas expresiones podemos ver que existen funciones reales $\boldsymbol{\alpha}(t)$ y $\boldsymbol{\beta}(t)$ que

$$\mathbf{f}''(t) = \boldsymbol{\alpha}(t)\mathbf{N}(s) + \boldsymbol{\beta}(t)\mathbf{T}(s)$$

(de hecho, $\boldsymbol{\alpha}(t) = \|\mathbf{f}'(t)\|^2$, $\boldsymbol{\beta}(t) = \frac{\mathbf{f}'(t)\cdot\mathbf{f}''(t)}{\|\mathbf{f}'(t)\|}$). Esto nos dice que el vector $\mathbf{f}''(t)$ se encuentra en el plano osculador de $\mathbf{f}$ en t, pues es una combinación lineal de los vectores $\mathbf{T}(s)$ y $\mathbf{N}(s)$, los cuales determinan dicho plano. Entonces el vector $\mathbf{v} = \mathbf{f}'(t) \times \mathbf{f}''(t)$ se encuentra en la dirección del vector binormal $B(s)$. Con este vector $\mathbf{v}$ podemos determinar el plano osculador y la recta binormal de $\mathbf{f}$ en t.

Resta ver lo correspondiente al plano rectificante y a la recta normal de $\mathbf{f}$ en t. Tenemos determinados ya los vectores $\mathbf{u} = \mathbf{f}'(t)$ y $\mathbf{v} = \mathbf{f}'(t) \times \mathbf{f}''(t)$ los cuales son ortogonales a los planos normal y osculador respectivamente. Entonces el vector $\mathbf{w} = (\mathbf{f}'(t) \times \mathbf{f}''(t)) \times \mathbf{f}'(t)$ es un vector ortogonal a ambos vectores $\mathbf{u}$ y $\mathbf{v}$ y por lo tanto, debe estar en la dirección del vector $\mathbf{N}(s)$. De hecho, $\mathbf{N}(s) = \frac{\mathbf{w}}{\|\mathbf{w}\|}$. Con este vector ya se determina de inmediato plano rectificante y la recta normal procurados.

En la figura 2 se muestran los vectores mencionados en la exposición anterior. Los números entre paréntesis indican el orden en que se deben ir obteniendo los vectores.

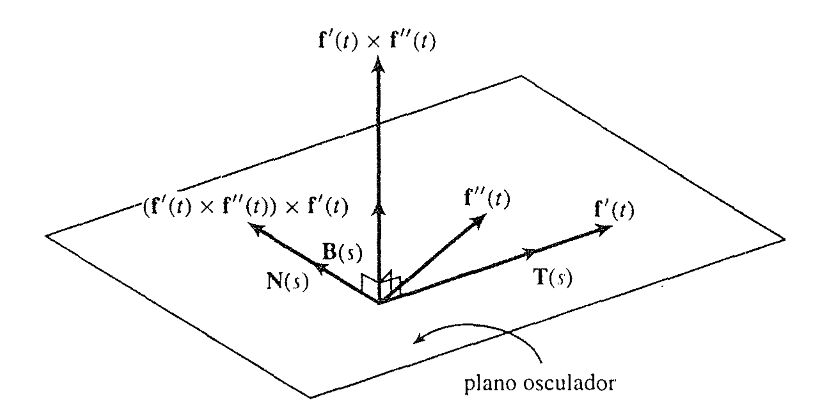


Figura 2.

Ejemplo 2. Considere el camino $\mathbf{f}(t) = (t, t^2, t^3)$. Se quieren determinar las ecuaciones de los planos osculador, normal y rectificante, así como las rectas tangente, normal y binormal a la curva

descrita por $\mathbf{f}$ en el punto $\mathbf{p} = \mathbf{f}(2) = (2, 4, 8)$. Tenemos

$$\mathbf{u} = \mathbf{f}'(t) = (1, 2t, 3t^2), \qquad \mathbf{f}''(t) = (0, 2, 6t)$$

$$\mathbf{v} = \mathbf{f}'(t) \times \mathbf{f}''(t) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2t & 3t^2 \\ 0 & 2 & 6t \end{bmatrix} = (6t^2, -6t, 2)$$

$$\mathbf{w} = \mathbf{v} \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 6t^2 & -6t & 2 \\ 1 & 2t & 3t^2 \end{bmatrix} = (-18t^3 - 4t, 2 - 18t^4, 12t^3 + 6t)$$

En $t = 2$, se tiene $\mathbf{u} = (1, 4, 12)$, $\mathbf{f}''(2) = (0, 2, 12)$, $\mathbf{v} = (24, -12, 2)$, $\mathbf{w} = (-152, -286, 108)$. El plano osculador pasa por $(2, 4, 8)$ y tiene a $\mathbf{v} = (24, -12, 2)$ por vector normal. Su ecuación es

$$24(x - 2) - 12(y - 4) + 2(z - 8) = 0$$

o bien

$$12x - 6y + z = 8$$

El plano normal pasa por $(2, 4, 8)$ y tiene a $\mathbf{u} = (1, 4, 12)$ por vector normal. Su ecuación es

$$1(x - 2) + 4(y - 4) + 12(z - 8) = 0$$

o bien

$$x + 4y + 12z = 114$$

El plano rectificante pasa por $(2, 4, 8)$ y tiene a $\mathbf{w} = (-152, -286, 108)$ por vector normal. Su ecuación es

$$-152(x - 2) - 286(y - 4) + 108(z - 8) = 0$$

o bien

$$76x + 143y - 54z = 292$$

Se obtienen fácilmente también las ecuaciones de las rectas tangentes, normal y binormal, que son

$$(x, y, z) = (2, 4, 8) + t(1, 4, 12)$$
$$(x, y, z) = (2, 4, 8) + t(-152, -286, 108)$$
$$(x, y, z) = (2, 4, 8) + (24, -12, 2), \qquad t \in \mathbb{R}$$

respectivamente. ■

Desde luego, los procedimientos descritos anteriormente para obtener las ecuaciones de los planos osculador, normal y rectificante, así como las de las rectas tangente, normal y binormal, no se alteran en absoluto por reparametrizaciones de algún camino dado.

Ejemplo 3. Se quiere hallar el plano osculador en la curva C que resulta de la intersección de la esfera $x^2 + y^2 + z^2 = 6$ con el paraboloide $z = x^2 + y^2$, en el punto $(1, 1, 2)$. Se puede verificar que el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ dado por $\mathbf{f}(t) = (\sqrt{2}\cos t, \sqrt{2}\operatorname{sen} t, 2)$ recorre la curva C, la cual es un

círculo de centro en el origen y radio $\sqrt{2}$ dibujado en el plano $z = 2$. Como vimos, el vector normal al plano osculador es $\mathbf{f}'(t) \times \mathbf{f}''(t)$. Se tiene

$$\mathbf{f}'(t) \times \mathbf{f}''(t) = (-\sqrt{2}\operatorname{sen} t, \sqrt{2}\cos t, 0) \times (-\sqrt{2}\cos t, -\sqrt{2}\operatorname{sen} t, 0)$$
$$= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sqrt{2}\operatorname{sen} t & \sqrt{2}\cos t & 0 \\ -\sqrt{2}\cos t & -\sqrt{2}\operatorname{sen} t & 0 \end{bmatrix} = (0, 0, 2)$$

Entonces el plano osculador procurado es

$$0(x - 1) + 0(y - 1) + 2(z - 2) = 0$$

o sea, $z = 2$. Nótese que en realidad este es el plano osculador del camino $\mathbf{f}$ en cualquier $t \in \mathbb{R}$, i.e. es el plano osculador del círculo C en cualquier punto de él.

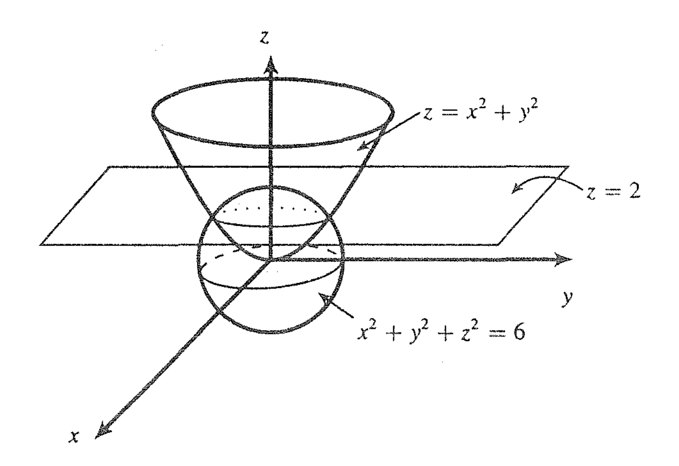


Figura 3. El plano osculador del ejemplo 3. ■

Ejemplo 4. Consideremos la curva C que resulta de la intersección del cilindro parabólico $y = x^2$ con el plano $z = 2x$. Se quieren obtener los vectores tangente unitario, normal principal, y binormal a la curva C en el punto $\mathbf{p} = (1, 1, 2)$. Un camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ que parametriza la curva C es $\mathbf{f}(t) = (t, t^2, 2t)$, siendo $\mathbf{p} = \mathbf{f}(1)$. El vector $\mathbf{f}'(t) = (1, 2t, 2)$ es el vector velocidad o tangente a la curva en $\mathbf{f}(t)$. Para $t = 1$ tenemos el vector $\mathbf{f}'(1) = (1, 2, 2)$, de modo que el vector tangente unitario es

$$\mathbf{T} = \frac{(1, 2, 2)}{\|(1, 2, 2)\|} = \frac{1}{3}(1, 2, 2)$$

El siguiente vector que hallaremos es el binormal, haciendo el producto cruz $\mathbf{f}'(t) \times \mathbf{f}''(t)$ en $t = 1$. Tenemos $\mathbf{f}''(t) = (0, 2, 0)$, de modo que

$$\mathbf{f}'(1) \times \mathbf{f}''(1) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 2 \\ 0 & 2 & 0 \end{bmatrix} = (-4, 0, 2)$$

El vector binormal lo obtenemos normalizando este vector

$$\mathbf{B} = \frac{(-4, 0, 2)}{\|(-4, 0, 2)\|} = \frac{1}{\sqrt{20}}(-4, 0, 2)$$

Por último el vector normal principal lo obtenemos como $\mathbf{N} = \mathbf{B} \times \mathbf{T}$

$$\mathbf{N} = \frac{1}{\sqrt{20}}(-4, 0, 2) \times \frac{1}{3}(1, 2, 2) = \frac{1}{3\sqrt{20}} \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -4 & 0 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

$$= \frac{1}{3\sqrt{20}}(-4, 10, -8) = \frac{1}{3\sqrt{5}}(-2, 5, -4) \qquad \blacksquare$$

Ejercicios (Capítulo 5, Sección 9)

En cada uno de los ejercicios 1–5, determine las ecuaciones de la recta tangente, la recta normal, la recta binormal, el plano osculador, el plano normal y el plano rectificante a la curva dada en el punto indicado.

1. $\boldsymbol{\lambda}(t) = (t, t^2, t^2 + 3)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(0)$.
2. $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t, \tan t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(0)$.
3. $\boldsymbol{\lambda}(t) = (e^t, e^{-t}, t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(0)$.
4. $\boldsymbol{\lambda}(t) = (t, \ln t, t^2)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(1)$.
5. $\boldsymbol{\lambda}(t) = (t^3, t^2, t)$, en el punto $\mathbf{p} = \boldsymbol{\lambda}(1)$.
6. Determine la ecuación del plano osculador a la curva (imagen del camino) $\boldsymbol{\lambda}: \mathbb{R} \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (a_1t^2 + b_1t + c_1, a_2t^2 + b_2t + c_2, a_3t^2 + b_3t + c_3)$, en un punto $\mathbf{p} = \boldsymbol{\lambda}(t)$ cualquiera. (Ver ejercicio 23 de la sección 2).
7. Demuestre que las rectas tangente, normal y binormal a la curva (imagen del camino) $\boldsymbol{\lambda}: \mathbb{R} \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (e^t \cos t, e^t \operatorname{sen} t, e^t)$ forman ángulos constantes con el eje z.
8. Determine las ecuaciones de los planos osculador, normal y rectificante a la curva de intersección de los cilindros $x^2 + y^2 = 25$, $z^2 + y^2 = 10$, en el punto $(4, 3, 1)$.
9. Determine las ecuaciones de los planos osculador, normal y rectificante a la curva de intersección de las superficies $x^2 + y^2 + z^2 = 3$, $x^2 + y^2 = 2$, en el punto $\mathbf{p} = (1, 1, 1)$.
10. Halle las ecuaciones de la rectas tangente, normal y binormal a la curva de intersección de las superficies $z = x^2$, $z = y^2$, en el punto $\mathbf{p} = (1, 1, 1)$.
11. Hallar el volumen del tetraedro que forma, con los planos coordenados, el plano osculador a la curva de intersección de las superficies $x^2 + y^2 + z^2 = 6$, $x + y + z = 4$, en el punto $(1, 1, 2)$. (Sugerencia: piense antes de comenzar a hacer operaciones).

5.10 Torsión

Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ un camino regular tres veces diferenciable parametrizado por longitud de arco. En la sección anterior vimos cómo obtener el plano osculador en cada punto de la curva que describe $\mathbf{f}$. Quisiéramos ahora centrar nuestra atención en la rapidez con que una curva se aleja de su plano osculador en la vecindad de un punto dado de ella. Esta rapidez está relacionada (directamente) con el concepto que estudiaremos en esta sección llamado torsión. Por ejemplo, viendo las dos figuras siguientes, pensamos (intuitivamente) que la curva C_2 tendrá "más torsión" en $\mathbf{p}$ que la curva C_1 en $\mathbf{p}$.

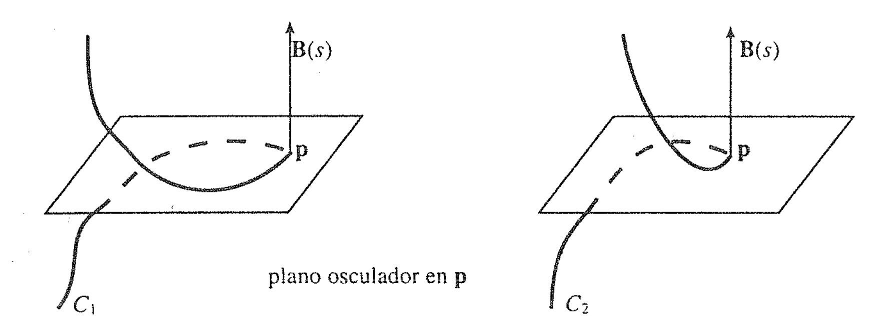


Figura 1. La curva C_2 tiene "más torsión" que la curva C_1 en p.

La forma de medir la rapidez de alejamiento de la curva C de su plano osculador es por medio del vector binormal $\mathbf{B}(s) = \mathbf{T}(s) \times \mathbf{N}(s)$, que sabemos es un vector unitario ortogonal al plano osculador de $\mathbf{f}$ en s. Puesto que $\|\mathbf{B}(s)\| = 1$, $\forall s \in I$, la magnitud de la derivada $\|\mathbf{B}(s)\|$ de $\mathbf{B}(s)$ medirá precisamente la rapidez con que el vector binormal $\mathbf{B}(s)$ está cambiando de dirección en los alrededores del punto estudiado.

Puesto que $\mathbf{B}(s) = \mathbf{T}(s) \times \mathbf{N}(s)$, tenemos, derivando

$$\mathbf{B}'(s) = \mathbf{T}(s) \times \mathbf{N}'(s) + \mathbf{T}'(s) \times \mathbf{N}(s)$$

El sumando $\mathbf{T}'(s) \times \mathbf{N}(s)$ que aparece en esta expresión es igual a cero, ya que $\mathbf{T}'(s) = \mathbf{f}''(s)$ es un vector en la dirección de $\mathbf{N}(s)$ (y son por lo tanto colineales; por lo que su producto cruz es cero). Entonces nos queda que

$$\mathbf{B}'(s) = \mathbf{T}(s) \times \mathbf{N}'(s)$$

Por otra parte, puesto que $\mathbf{B}(s)$ es un vector unitario, tenemos que el vector $\mathbf{B}'(s)$ es ortogonal a $\mathbf{B}(s)$ (de $\|\mathbf{B}(s)\| = 1$ se deduce, derivando que $\mathbf{B}(s) \cdot \mathbf{B}'(s) = 0$).

En resumen, tenemos los hechos siguientes:

i. $\mathbf{B}'(s)$ es un vector ortogonal a $\mathbf{B}(s)$. Esto nos permite concluir que $\mathbf{B}'(s)$ es un vector en el plano osculador de $\mathbf{f}$ en s.

ii. $\mathbf{B}'(s) = \mathbf{T}(s) \times \mathbf{N}'(s)$, lo cual nos permite concluir que $\mathbf{B}'(s)$ es un vector ortogonal a $\mathbf{T}(s)$.

Conclusión: $\mathbf{B}'(s)$ es un vector paralelo al vector $\mathbf{N}(s)$. Debe haber entonces un escalar $\tau(s)$ que

$$\mathbf{B}'(s) = \tau(s)\mathbf{N}(s)$$

Nótese que $|\tau(s)| = \|\mathbf{B}'(s)\|$.

Por supuesto, la discusión anterior tiene sentido solamente en caso de que la curvatura de $\mathbf{f}$ en s es no nula, pues es justamente cuando podemos hablar del plano osculador de $\mathbf{f}$ en s, que fue el punto de arranque de la discusión (si en un punto la curvatura es cero, el vector $\mathbf{N}(s)$ no queda bien determinado, y en consecuencia, el plano osculador de $\mathbf{f}$ en s queda también sin determinarse). Establezcamos formalmente la definición de torsión.

Definición. Sea $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ un camino regular tres veces diferenciable parametrizado por longitud de arco tal que $\mathbf{f}''(s) \neq 0\ \forall s \in I$ (i.e. la curvatura $k(s)$ es siempre no nula). El número real $\tau(s)$ tal que $\mathbf{B}'(s) = \tau(s)\mathbf{N}(s)$, donde $\mathbf{B}(s)$ y $\mathbf{N}(s)$ son los vectores binormal y normal principal de $\mathbf{f}$ en s respectivamente, se llama *torsión* de $\mathbf{f}$ en s. ■

Así como se mencionó que la idea intuitiva de la curvatura de una curva era "la medida de cuánto se 'curva' la curva", la idea intuitiva análoga del concepto de torsión es "la medida de cuánto se 'tuerce' la curva". De hecho, una curva en el espacio queda caracterizada por completo por estas dos medidas: la rapidez con que se 'curva' y la rapidez con que se 'tuerce'.

A la luz de estas ideas, veamos que una curva es *plana* (es decir, es la imagen de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ tal que $\mathbf{f}(s)$ se encuentra en un plano en $\mathbb{R}^3$ para toda $s \in I$) si y sólo si su torsión es igual a cero ($\forall s \in I$). En efecto, si $\tau(s) = 0$, se tiene $\mathbf{B}'(s) = 0$, por lo que el vector binormal $\mathbf{B}(s)$ debe ser constante, i.e. $\mathbf{B}(s) = \mathbf{v}\ \forall s \in I$. De aquí se tiene que $\mathbf{f}'(s) \cdot \mathbf{B}(s) = \mathbf{f}'(s) \cdot \mathbf{v} = 0\ \forall s \in I$ (pues $\mathbf{f}'(s)$ es ortogonal a $\mathbf{B}(s)$), o sea

$$\frac{d}{ds}\left(\mathbf{f}(s) \cdot \mathbf{v}\right) = 0$$

de donde $\mathbf{f}(s) \cdot \mathbf{v} = \text{cte}\ \forall s \in I$, y por lo tanto, concluimos finalmente que $\mathbf{f}(s)$ se encuentra en el plano cuyo vector normal es $\mathbf{v}$, $\forall s \in I$ (figura 2).

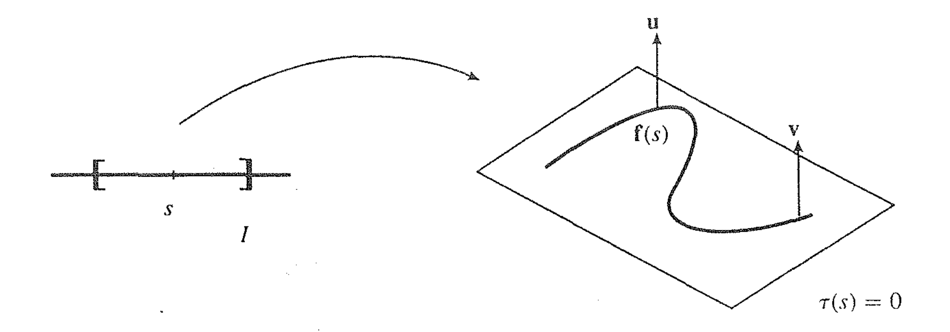


Figura 2. Una curva plana.

Recíprocamente, si $\mathbf{f}(s)$ se encuentra en un plano para toda s en I, entonces dicho plano es el plano osculador de la curva en todo punto de ella, por tanto el vector unitario $\mathbf{B}(s)$ no cambia de dirección, por lo que $\mathbf{B}'(s) = 0$, de donde $\tau(s) = 0\ \forall s \in I$.

Como suele suceder con algunas definiciones en matemáticas, ésta no proporciona una manera práctica de calcular los conceptos definidos (así ocurrió con la curvatura). Si bien la definición

de torsión dada anteriormente nos permite ver con claridad qué es lo que mide la torsión, no nos dice cómo calcular prácticamente la torsión de un camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^3$. A continuación nos proponemos obtener una fórmula que nos permita calcular la torsión de un camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^3$ regular, tres veces diferenciable, con curvatura no nula, el cual no está (necesariamente) parametrizado por longitud de arco.

Sea $\bar{\mathbf{f}}\colon J \subseteq \mathbb{R} \to \mathbb{R}^3$ la reparametrización por longitud de arco de $\mathbf{f}$, $\bar{\mathbf{f}} = \mathbf{f} \circ \varphi$. Queremos calcular la torsión de $\mathbf{f}$ en t, donde $t = \varphi(s)$. Sabemos que

$$\begin{aligned}
\mathbf{T}(s) &= \bar{\mathbf{f}}'(s) = \frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t),\\
\mathbf{T}'(s) &= \bar{\mathbf{f}}''(s) = \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - (\mathbf{f}'(t)\cdot\mathbf{f}''(t))\mathbf{f}'(t)\right)\\
\mathbf{N}(s) &= \frac{1}{k(s)}\bar{\mathbf{f}}''(s) = \frac{1}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}''(s)
\end{aligned}$$

Calculemos $\mathbf{B}'(s) = \mathbf{T}(s) \times \mathbf{N}'(s)$. Se tiene

$$\mathbf{N}'(s) = \frac{\|\bar{\mathbf{f}}''(s)\|\bar{\mathbf{f}}'''(s) - \dfrac{\bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}'''(s)}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}''(s)}{\|\bar{\mathbf{f}}''(s)\|^2} = \frac{1}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}'''(s) - \frac{\bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}'''(s)}{\|\bar{\mathbf{f}}''(s)\|^3}\bar{\mathbf{f}}''(s)$$

Entonces

$$\begin{aligned}
\mathbf{B}'(s) &= \mathbf{T}(s) \times \mathbf{N}'(s)\\
&= \bar{\mathbf{f}}'(s) \times \left(\frac{1}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}'''(s) - \frac{\bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}'''(s)}{\|\bar{\mathbf{f}}''(s)\|^2}\bar{\mathbf{f}}''(s)\right)\\
&= \frac{1}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}'''(s) - \frac{\bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}'''(s)}{\|\bar{\mathbf{f}}''(s)\|^3}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s)
\end{aligned}$$

La torsión $\tau(s)$ está determinada por $\mathbf{B}'(s) = \tau(s)\mathbf{N}(s)$. Al tomar el producto punto con el vector $\mathbf{N}(s)$ en ambos miembros de la última expresión nos queda

$$\mathbf{B}'(s)\cdot\mathbf{N}(s) = \tau(s)\mathbf{N}(s)\cdot\mathbf{N}(s) = \tau(s)\|\mathbf{N}(s)\|^2 = \tau(s)$$

Entonces

$$\begin{aligned}
\tau(s) &= \mathbf{B}'(s)\cdot\mathbf{N}(s)\\
&= \left(\frac{1}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}'''(s) - \frac{\bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}'''(s)}{\|\bar{\mathbf{f}}''(s)\|^3}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s)\right)\cdot\frac{1}{\|\bar{\mathbf{f}}''(s)\|}\bar{\mathbf{f}}''(s)\\
&= \frac{1}{\|\bar{\mathbf{f}}''(s)\|^2}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}'''(s)\cdot\bar{\mathbf{f}}''(s) - \frac{\bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}'''(s)}{\|\bar{\mathbf{f}}''(s)\|^4}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s)\cdot\bar{\mathbf{f}}''(s)\\
&= \frac{1}{\|\bar{\mathbf{f}}''(s)\|^2}\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}'''(s)\cdot\bar{\mathbf{f}}''(s)
\end{aligned}$$

(pues $\bar{\mathbf{f}}''(s)$ es ortogonal a $\bar{\mathbf{f}}'(s)\times\bar{\mathbf{f}}''(s)$ y por lo tanto su producto punto es cero). Como $k(s) = \|\bar{\mathbf{f}}''(s)\|$, tenemos la fórmula

$$\tau(s) = \frac{\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}'''(s)\cdot\bar{\mathbf{f}}''(s)}{(k(s))^2}$$

Esta fórmula nos permite calcular la torsión de un camino $\bar{\mathbf{f}}$ (regular tres veces diferenciable) el cual está parametrizado por una longitud de arco. Para quitar esta restricción, escribamos las derivadas $\bar{\mathbf{f}}'(s)$, $\bar{\mathbf{f}}''(s)$ y $\bar{\mathbf{f}}'''(s)$ del camino $\bar{\mathbf{f}}$ en términos del camino $\mathbf{f}$ dado originalmente. Obsérvese que

$$\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}'''(s) \cdot \bar{\mathbf{f}}''(s) = -\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s) \cdot \bar{\mathbf{f}}'''(s)$$

Hagamos explícito el vector $\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s)$

$$\begin{aligned}
\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s) &= \frac{1}{\|\mathbf{f}'(t)\|}\mathbf{f}'(t) \times \frac{1}{\|\mathbf{f}'(t)\|^4}\left(\|\mathbf{f}'(t)\|^2\mathbf{f}''(t) - (\mathbf{f}'(t) \cdot \mathbf{f}''(t))\mathbf{f}'(t)\right) \\
&= \frac{1}{\|\mathbf{f}'(t)\|^3}\mathbf{f}'(t) \times \mathbf{f}''(t) - \frac{\mathbf{f}'(t) \cdot \mathbf{f}''(t)}{\|\mathbf{f}'(t)\|^5}\mathbf{f}'(t) \times \mathbf{f}'(t) \\
&= \frac{1}{\|\mathbf{f}'(t)\|^3}\mathbf{f}'(t) \times \mathbf{f}''(t)
\end{aligned}$$

Obtengamos ahora el vector $\bar{\mathbf{f}}'''(s)$ (en términos de $\mathbf{f}$). Se tiene

$$\begin{aligned}
\bar{\mathbf{f}}'''(s) &= \frac{d}{ds}\bar{\mathbf{f}}''(s) = \left(\frac{d}{dt}\bar{\mathbf{f}}''(s)\right)\frac{dt}{ds} = \frac{1}{\|\mathbf{f}'(t)\|}\frac{d}{dt}\bar{\mathbf{f}}''(s) \\
&= \frac{1}{\|\mathbf{f}'(t)\|}\frac{d}{dt}\left(\frac{1}{\|\mathbf{f}'(t)\|^2}\mathbf{f}''(t) - \frac{\mathbf{f}'(t) \cdot \mathbf{f}''(t)}{\|\mathbf{f}'(t)\|^4}\mathbf{f}'(t)\right) \\
&= \frac{1}{\|\mathbf{f}'(t)\|}\left[\frac{1}{\|\mathbf{f}'(t)\|^2}\mathbf{f}'''(t) + \left(\frac{1}{\|\mathbf{f}'(t)\|^2}\right)'\mathbf{f}''(t)\right. \\
&\qquad \left. - \frac{\mathbf{f}'(t) \cdot \mathbf{f}''(t)}{\|\mathbf{f}'(t)\|^4}\mathbf{f}''(t) - \left(\frac{\mathbf{f}'(t) \cdot \mathbf{f}''(t)}{\|\mathbf{f}'(t)\|^4}\right)'\mathbf{f}'(t)\right]
\end{aligned}$$

donde la coma $'$ indica derivación respecto de t.

Téngase presente que al tomar el producto punto de este vector $\bar{\mathbf{f}}''(s)$ con el vector

$$\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s) = \frac{1}{\|\mathbf{f}'(t)\|^3}\mathbf{f}'(t) \times \mathbf{f}''(t)$$

y al aplicar la linealidad del producto punto, obtenemos tres sumandos iguales a cero, a saber, los correspondientes a los últimos tres sumandos que aparecen en la expresión de $\bar{\mathbf{f}}'''(s)$, en los que se involucran los vectores $\mathbf{f}''(t)$ y $\mathbf{f}'(t)$, que son ortogonales a $\mathbf{f}'(t) \times \mathbf{f}''(t)$. Nos queda entonces que

$$\begin{aligned}
\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s) \cdot \bar{\mathbf{f}}'''(s) &= \left(\frac{1}{\|\mathbf{f}'(t)\|^3}\mathbf{f}'(t) \times \mathbf{f}''(t)\right) \cdot \left(\frac{1}{\|\mathbf{f}'(t)\|^3}\mathbf{f}'''(t)\right) \\
&= \frac{1}{\|\mathbf{f}'(t)\|^6}\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t)
\end{aligned}$$

Regresando a la fórmula de la torsión y haciendo explícita la curvatura $k(s)$ en términos del camino $\mathbf{f}$ (ver página 489) nos queda

$$\tau(t) = -\frac{\bar{\mathbf{f}}'(s) \times \bar{\mathbf{f}}''(s) \cdot \bar{\mathbf{f}}'''(s)}{(k(s))^2} = -\frac{\frac{1}{\|\mathbf{f}'(t)\|^6}\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t)}{\left(\frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3}\right)^2}$$

de la que resulta finalmente la fórmula

$$\boxed{\tau(t) = -\frac{\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t)}{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|^2}}$$

que nos da el valor de la torsión del camino $\mathbf{f}$ en t.

Ejemplo 1. Sea $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^3$ el camino $\mathbf{f}(t) = (\alpha \cos t, \alpha \operatorname{sen} t, \beta t)$. Este es un camino regular tres veces diferenciable con curvatura no nula (de hecho $k(t) = \frac{|\alpha|}{\alpha^2+\beta^2}$ $\forall t \in \mathbb{R}$, ver ejemplo 4 de la sección 7). La curva que $\mathbf{f}$ describe es una hélice circular de curvatura constante. Calculemos su torsión. Se tiene

$$\mathbf{f}'(t) = (-\alpha \operatorname{sen} t, \alpha \cos t, \beta)$$
$$\mathbf{f}''(t) = (-\alpha \cos t, -\alpha \operatorname{sen} t, 0)$$
$$\mathbf{f}'''(t) = (\alpha \operatorname{sen} t, -\alpha \cos t, 0)$$
$$\mathbf{f}'(t) \times \mathbf{f}''(t) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\alpha \operatorname{sen} t & \alpha \cos t & \beta \\ -\alpha \cos t & -\alpha \operatorname{sen} t & 0 \end{bmatrix} = (\alpha\beta \operatorname{sen} t, -\alpha\beta \cos t, \alpha^2)$$
$$\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t) = (\alpha\beta \operatorname{sen} t, -\alpha\beta \cos t, \alpha^2) \cdot (\alpha \operatorname{sen} t, -\alpha \cos t, 0) = \alpha^2\beta$$
$$\|\mathbf{f}'(t) \times \mathbf{f}''(t)\| = \|(\alpha\beta \operatorname{sen} t, -\alpha\beta \cos t, \alpha^2)\| = \sqrt{\alpha^2\beta^2 + \alpha^4} = \alpha\sqrt{\alpha^2 + \beta^2}$$

Entonces

$$\tau(t) = -\frac{\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t)}{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|^2} = -\frac{\alpha^2\beta}{\left(\alpha\sqrt{\alpha^2+\beta^2}\right)^2} = -\frac{\beta}{\alpha^2 + \beta^2}$$

Así pues, la hélice que describe $\mathbf{f}$ tiene, además de la curvatura constante, la torsión constante. Obsérvese que $\tau(t) = 0$ si y sólo si $\beta = 0$. En tal caso el camino $\mathbf{f}$ es $\mathbf{f}(t) = (\alpha \cos t, \alpha \operatorname{sen} t, 0)$, el cual describe un círculo de centro en el origen y radio α en el plano xy (es una curva plana, como tenía que ocurrir). ■

Ejemplo 2. Considere el camino del ejemplo 5 de la sección 7, $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (2 + t, 1 + t^2, 3t + t^2)$. La curvatura de $\mathbf{f}$ en t (como se calculó en el ejemplo mencionado)

$$k(t) = \frac{\sqrt{44}}{(10 + 12t + 8t^2)^{3/2}}, \quad \text{(no es constante)}$$

Calculemos su torsión. Se tiene

$$\mathbf{f}'(t) = (1, 2t, 3 + 2t)$$
$$\mathbf{f}''(t) = (0, 2, 2)$$
$$\mathbf{f}'''(t) = (0, 0, 0)$$

Entonces

$$\tau(t) = \frac{\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t)}{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|^2} = 0$$

Es decir, la curva que describe el camino $\mathbf{f}$ es plana. Veamos en qué plano se encuentra esta curva. Este plano debe ser el plano osculador de la curva (en cualquier punto de ella), el cual, sabemos que tiene por vector normal al vector binormal $\mathbf{B}(t)$, o bien, como se vió en la sección anterior, al vector $\mathbf{f}'(t) \times \mathbf{f}''(t)$, que es

$$\mathbf{f}'(t) \times \mathbf{f}''(t) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2t & 3 + 2t \\ 0 & 2 & 2 \end{bmatrix} = (-6, -2, 2)$$

(Nótese que este vector es el mismo en todos los puntos de la curva, como tenía que ocurrir). Entonces el plano osculador de la curva descrita por $\mathbf{f}$ en el punto $\mathbf{f}(t) = (2 + t, 1 + t^2, 3t + t^2)$ es

$$-6\big(x - (2 + t)\big) - 2\big(y - (1 + t^2)\big) + 2\big(z - (3t + t^2)\big) = 0$$

o bien

$$-6x - 2y - 2z + 6(2 + t) + 2(1 + t^2) - 2(3t + t^2) = 0$$

o, finalmente

$$3x + y - z - 7 = 0$$

que no depende del punto $\mathbf{f}(t)$ del camino $\mathbf{f}$. ■

Ejemplo 3. Sea $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$ el camino $\mathbf{f}(t) = (\sqrt{2}\cos t, \sqrt{2}\,\mathrm{sen}\, t, 2)$. En el ejemplo 3 de la sección anterior vimos que este camino describe el círculo resultante de la intersección de la esfera $x^2 + y^2 + z^2 = 6$ con el paraboloide $z = x^2 + y^2$. Luego se trata de una curva plana. Verifiquemos este hecho calculando su torsión. Se tiene

$$\mathbf{f}'(t) = (-\sqrt{2}\,\mathrm{sen}\, t, \sqrt{2}\cos t, 0)$$
$$\mathbf{f}''(t) = (-\sqrt{2}\cos t, -\sqrt{2}\,\mathrm{sen}\, t, 0)$$
$$\mathbf{f}'''(t) = (\sqrt{2}\,\mathrm{sen}\, t, -\sqrt{2}\cos t, 0)$$
$$\mathbf{f}'(t) \times \mathbf{f}''(t) = (0, 0, 2)$$
$$\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t) = 0(\sqrt{2}\,\mathrm{sen}\, t) + 0(-\sqrt{2}\cos t) + 2(0) = 0$$

de modo que

$$\tau(t) = -\frac{\mathbf{f}'(t) \times \mathbf{f}''(t) \cdot \mathbf{f}'''(t)}{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|^2} = 0$$

para todo $t \in \mathbb{R}$, como tenía que ocurrir. ■

Así como aconteció con la curvatura, la torsión de un camino **f** en t es invariante por reparametrizaciones de **f**, de modo que, efectivamente, la torsión es una propiedad geométrica de la curva que describe **f** y no depende de cómo sea recorrida la curva. Queda la demostración de este hecho como un ejercicio para el lector.

Ejemplo 4. Se quiere hallar la torsión de la curva C que resulta de la intersección de la superficie $z = 2x^2y$ con la superficie $z = x + y$ en el punto $(1, 1, 2)$. Tomando como parámetro $x = t$, tenemos (resolviendo simultáneamente las ecuaciones que describe a las superficies) $2x^2y = x + y$, de donde $y = \frac{x}{2x^2-1} = \frac{t}{2t^2-1}$ y finalmente $z = 2x^2y = \frac{2t^3}{2t^2-1}$. Entonces el camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^3$ dado por

$$\mathbf{f}(t) = \left(t, \frac{t}{2t^2-1}, \frac{2t^3}{2t^2-1}\right)$$

describe la curva C, al menos en los alrededores del punto $\mathbf{p} = \mathbf{f}(1) = (1, 1, 2)$. Se tiene

$$\mathbf{f}'(t) = \left(1, -\frac{2t^2+1}{(2t^2-1)^2}, \frac{4t^4-6t^2}{(2t^2-1)^2}\right)$$

$$\mathbf{f}''(t) = \left(0, \frac{8t^3+12t}{(2t^2-1)^3}, \frac{8t^3+12t}{(2t^2-1)^3}\right)$$

$$\mathbf{f}'''(t) = \left(0, -\frac{48t^4+144t^2+12}{(2t^2-1)^4}, -\frac{48t^4+144t^2+12}{(2t^2-1)^4}\right)$$

de modo que en $t = 1$, los vectores son

$$\mathbf{f}'(1) = (1, -3, -2), \quad \mathbf{f}''(1) = (0, 20, 20), \quad \mathbf{f}'''(1) = (0, -204, -204)$$

Entonces, como

$$\mathbf{f}'(1) \times \mathbf{f}''(1) \cdot \mathbf{f}'''(1) = \det \begin{bmatrix} 1 & -3 & -2 \\ 0 & 20 & 20 \\ 0 & -204 & -204 \end{bmatrix} = 0$$

La torsión de la curva C en $(1, 1, 2)$ es cero. ■

Se ha visto ya cómo en cada punto de la curva que describe un camino $\mathbf{f}\colon I \subseteq \mathbb{R} \to \mathbb{R}^3$ regular tres veces diferenciable con curvatura no nula se pueden asociar los vectores ortonormales: el vector tangente unitario **T**, el vector normal principal **N**, y el vector binormal **B**. Estos tres vectores forman, en cada punto **p** de la curva, un sistema ortonormal respecto del cual se puede estudiar el comportamiento local de la curva en los alrededores del punto **p**. Esta idea es una de las más importantes en el estudio de la geometría de las curvas.

Quisiéramos terminar esta sección estableciendo unas fórmulas célebres que relacionan las derivadas de los vectores **T**, **N** y **B** en términos del sistema de mencionado.

Supongamos que el camino **f** está parametrizado por longitud de arco. Entonces las definiciones de curvatura y de torsión dicen que

$$\mathbf{T}' = k\mathbf{N}, \quad \mathbf{B}' = \tau\mathbf{N}$$

fórmulas que nos dan las derivadas de **T** y **B** respecto de los vectores **T**, **N** y **B**. Veamos cómo expresar la derivada **N**′ de **N** (respecto de s, como la de **T** y **B** anteriormente) en términos de los vectores del sistema ortonormal **T**, **N** y **B**.

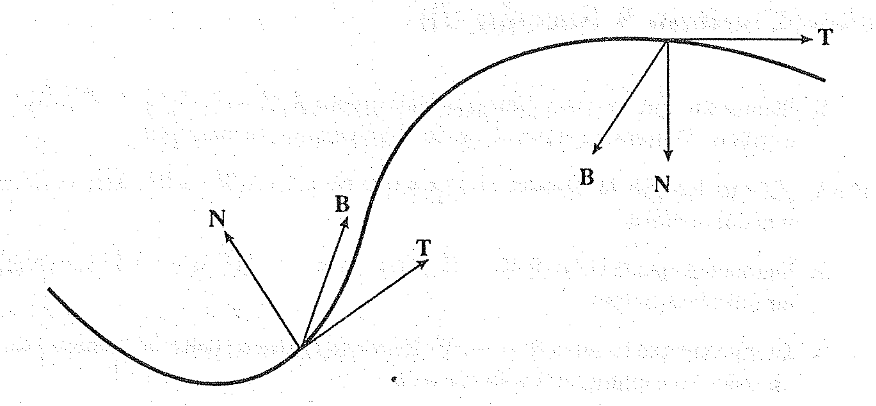


Figura 3. Sistema de referencia que forman **T**, **N** y **B**.

Tomemos como base el hecho de que si el conjunto $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortonormal de vectores de $\mathbb{R}^3$ entonces el vector $\mathbf{v} \in \mathbb{R}^3$ se escribe en términos de los vectores $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$ como

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v} \cdot \mathbf{u}_3)\mathbf{u}_3$$

Entonces, en nuestro caso

$$\mathbf{N}' = (\mathbf{N}' \cdot \mathbf{T})\mathbf{T} + (\mathbf{N}' \cdot \mathbf{N})\mathbf{N} + (\mathbf{N}' \cdot \mathbf{B})\mathbf{B}$$

Ahora bien, como $\|\mathbf{N}\| = 1$, se deduce que $\mathbf{N}' \cdot \mathbf{N} = 0$. De la expresión $\mathbf{N} \cdot \mathbf{T} = 0$ (la cual establece la ortogonalidad de los vectores **N** y **T**) se obtiene derivando:

$$\mathbf{N}' \cdot \mathbf{T} + \mathbf{N} \cdot \mathbf{T}' = 0$$

de donde (usando que $\mathbf{T}' = k\mathbf{N}$)

$$\mathbf{N}' \cdot \mathbf{T} = -\mathbf{N} \cdot \mathbf{T}' = -\mathbf{N} \cdot (k\mathbf{N}) = -k\|\mathbf{N}\|^2 = -k$$

del mismo modo, de $\mathbf{N} \cdot \mathbf{B} = 0$ se obtiene, derivando

$$\mathbf{N}' \cdot \mathbf{B} + \mathbf{N} \cdot \mathbf{B}' = 0$$

de donde (Con $\mathbf{B}' = \tau\mathbf{N}$)

$$\mathbf{N}' \cdot \mathbf{B} = -\mathbf{N} \cdot \mathbf{B}' = -\mathbf{N} \cdot (\tau N) = -\tau\|\mathbf{N}\|^2 = -\tau$$

Finalmente se llegó a

$$\mathbf{N}' = -k\mathbf{T} - \tau\mathbf{B}$$

A las fórmulas

$$\boxed{\mathbf{T}' = k\mathbf{N}, \quad \mathbf{N}' = -k\mathbf{T} - \tau\mathbf{B}, \quad \mathbf{B}' = \tau\mathbf{N}}$$

se les conoce como "Fórmulas de Frenet". Estas constituyen una de las herramientas fundamentales en el estudio de la geometría (diferencial) de las curvas en el espacio.

Ejercicios (Capítulo 5, Sección 10)

1. Demuestre que la curva (imagen del camino) $\boldsymbol{\lambda}: \mathbb{R} - \{-1, 1\} \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = \left(\frac{1+t}{1-t}, \frac{1}{1-t^2}, \frac{1}{1+t}\right)$, es plana. Determine la ecuación del plano en que se encuentra

(*) 2. ¿Cómo debe ser la función $f(t)$ para que la curva $\boldsymbol{\lambda}: \mathbb{R} \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (a \cos t, a \operatorname{sen} t, f(t))$ sea una curva plana?

3. Demuestre que la curva $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (3t - t^3, 3t^2, 3t + t^3)$ tiene curvatura y torsión iguales en todos sus puntos.

4. Demuestre que la curva $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (a \cosh t, a \operatorname{senh} t, bt)$ tiene curvatura y torsión iguales en todos sus puntos, si y sólo si $a = b$.

5. Considere la hélice $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (a \cos t, a \operatorname{sen} t, bt)$. Demuestre que:
   a. el ángulo que forma la recta tangente con el eje z es constante.
   b. la recta normal en cualquier punto de la curva es perpendicular al eje z.
   c. el ángulo que forma la recta binormal con el eje z es constante.
   d. el cociente de la curvatura entre la torsión es constante.

6. Considere la curva $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t \cos t, t \operatorname{sen} t, t)$. Demuestre que las funciones coordenadas $x = x(t)$, $y = y(t)$, $z = z(t)$ de la función $\mathbf{f}$ satisfacen $z^2 = x^2 + y^2$ (que es la ecuación de un cono).
   a. Demuestre que el ángulo que forma la recta tangente con el eje z es constante.
   b. Demuestre que la recta normal en cualquier punto de la curva es perpendicular al eje z.
   c. Demuestre que el ángulo que forma la recta binormal con el eje z es constante.
   d. Calcule la curvatura de $\mathbf{f}$.
   e. Calcule la torsión de $\mathbf{f}$.
   f. Compruebe que el cociente de la curvatura entre la torsión es constante.
   (A la curva, imagen del camino $\mathbf{f}$, se le llama *hélice cónica*).

(*) 7. Se llama *hélice generalizada* a la imagen de un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ tal que su vector tangente $\mathbf{f}'(t)$ forma siempre un ángulo constante con un vector $\mathbf{v} \in \mathbb{R}^3$ dado. Demuestre que la curva, imagen del camino $\mathbf{f}$, es una hélice generalizada, si y sólo si:
   a. el ángulo que forma la recta tangente con el eje z es constante.
   b. la recta normal en cualquier punto de la curva es perpendicular al eje z.
   c. el ángulo que forma la recta binormal con el eje z es constante.
   d. el cociente de la curvatura entre la torsión es constante.

(*) 8. Demuestre que la curva $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^3$ es una hélice generalizada si y sólo si $\mathbf{f}^{(2)}(t) \times \mathbf{f}^{(3)}(t) \cdot \mathbf{f}^{(4)}(t) = 0$.

9. Demuestre que la curva de intersección de las superficies $x^2 = 3y$, $2xy = 9z$ es una hélice generalizada.

10. Demuestre que la imagen del camino $\mathbf{f}: \mathbb{R}^+ \to \mathbb{R}^3$, $\mathbf{f}(t) = (2t, \ln t, t^2)$ es una hélice generalizada.

5.11 Aplicaciones a la dinámica

En esta última sección del presente capítulo veremos algunas de las importantes aplicaciones de la teoría expuesta en la descripción de movimientos de cuerpos en el plano y en el espacio. La mayoría de los aspectos matemáticos aquí usados ya se estudiaron en alguna o algunas de las secciones anteriores (excepto la formulación de las ecuaciones de movimiento en coordenadas polares), así que la novedad en esta sección será principalmente el contenido físico con que se analizarán algunas fórmulas, además de nuevas notaciones para algunas de las funciones que hemos manejado en este capítulo.

Por supuesto que para un camino $\mathbf{f}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ o $\mathbb{R}^3$, pensaremos en su dominio I como un intervalo de tiempo y su traza como "la carretera de sus imágenes" que describe el punto $\mathbf{f}(t) \in \mathbb{R}^2$ o $\mathbb{R}^3$. Es usual, en lugar de la letra $\mathbf{f}$, usar la letra $\mathbf{r}$ para denotar la función. Más aún, al vector $\mathbf{r}(t)$ se le suele llamar "vector posición", el cual nos da la posición del punto que se mueve (según $\mathbf{r}$) en el instante t.

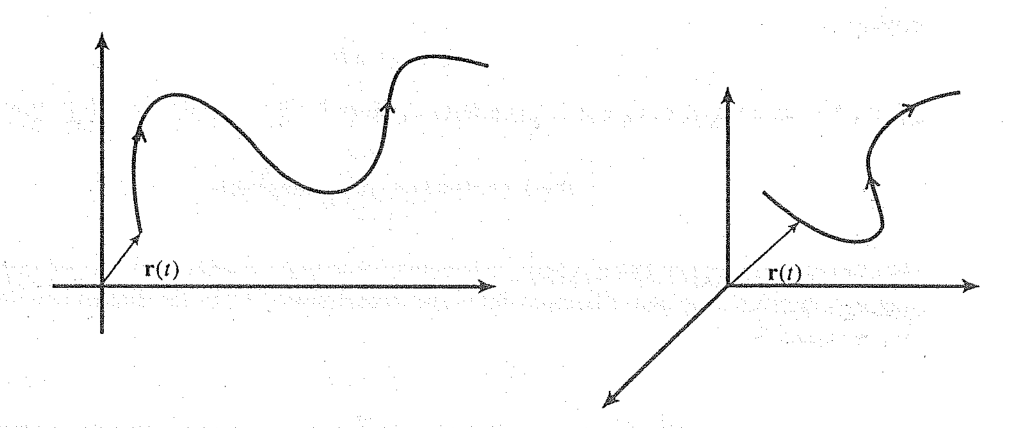


Figura 1. Gráfica que muestra al vector $\mathbf{r}(t)$.

Como el objetivo de esta sección es el contenido físico de las exposiciones presentadas hasta aquí, supondremos que las funciones $\mathbf{r}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ o $\mathbb{R}^3$ que aparezcan cumplen con todas las hipótesis de diferenciabilidad, regularidad y, en su caso, de curvatura no nula, necesarias para que las afirmaciones hechas sobre ellas tengan sentido.

Si escribimos $\mathbf{r}(t) = (x(t), y(t), z(t))$, al vector $\mathbf{r}'(t) = (x'(t), y'(t), z'(t))$ se le llama *vector velocidad*, y al vector $\mathbf{r}''(t) = (x''(t), y''(t), z''(t))$ se le llama *vector aceleración*, del camino o trayectoria $\mathbf{r}(t)$ en el instante t. De esta manera, la función $\mathbf{r}(t)$ nos da la información sobre el *movimiento* de un punto en el espacio, a saber, $\mathbf{r}(t)$ nos da la posición, $\mathbf{r}'(t)$ la velocidad y $\mathbf{r}''(t)$ la aceleración del movimiento en el instante t.

Seguimos llamando a $\|\mathbf{r}'(t)\|$ la *rapidez* del movimiento en el instante t, a $\mathbf{T}(t)$ el *vector tangente unitario* y a $\mathbf{N} = \mathbf{N}(t)$ el *vector normal principal* en t. Es decir, $\mathbf{T}(t) = \frac{1}{\|\mathbf{r}'(t)\|}\mathbf{r}'(t)$, y $\mathbf{N}(t)$ es el vector unitario ortogonal a $\mathbf{T} = \mathbf{T}(t)$, que, junto con éste último, determinan el plano osculador de $\mathbf{r}$ en t.

Escribamos $v(t) = \|\mathbf{r}'(t)\|$. Tenemos entonces que

$$\mathbf{r}'(t) = \|\mathbf{r}'(t)\|\mathbf{T}(t) = v(t)\mathbf{T}(t)$$

Derivando esta expresión respecto de t obtenemos

$$\mathbf{r}''(t) = v'(t)\mathbf{T}(t) + v(t)\mathbf{T}'(t)$$

Ahora bien, sabemos que la derivada del vector tangente unitário $\mathbf{T}$ *respecto de la longitud de arco* s, es igual, según la primera fórmula de Frenet (o bien, según la definición de curvatura) la curvatura k multiplicada por el vector normal principal $\mathbf{N}$, es decir, $\mathbf{T}'_s = k\mathbf{N}$, donde el subíndice s indica que la derivada está tomada respecto de s. Según la regla de la cadena tenemos

$$\mathbf{T}'(t) = \mathbf{T}'_t = \left(\mathbf{T}'_s\right)\frac{ds}{dt} = k\mathbf{N}\frac{ds}{dt}$$

Pero,

$$\frac{ds}{dt} = \frac{d}{dt}\left(\int_{t_0}^{t} \|\mathbf{r}'(u)\|du\right) = \|\mathbf{r}'(t)\| = v(t)$$

entonces

$$\mathbf{T}'(t) = kv\mathbf{N}$$

(k, v y $\mathbf{N}$ evaluados en t), de modo que sustituyendo en la expresión de $\mathbf{r}''(t)$ llegamos finalmente a

$$\boxed{\mathbf{r}''(t) = v''(t)\mathbf{T}(t) + k(t)(v(t))^2\mathbf{N}(t)}$$

Esta bonita e interesante fórmula (que ya habíamos obtenido y usando en varias ocasiones ¿dónde?) nos da la descomposición (ortogonal) del vector aceleración $\mathbf{r}''(t)$ en las direcciones de los vectores $\mathbf{T}$ y $\mathbf{N}$ (figura 2).

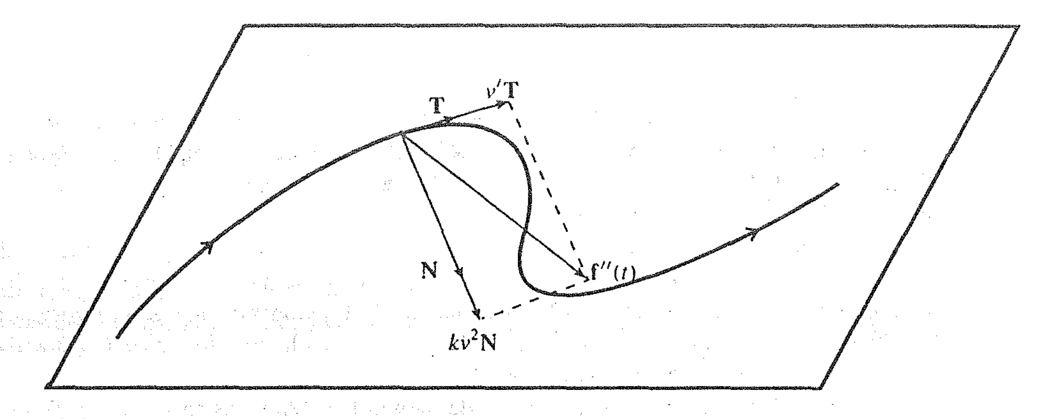


Figura 2. La descomposición de $\mathbf{r}''(t)$ en las direcciones de $\mathbf{N}$ y $\mathbf{T}$.

Observamos que la componente tangencial de $\mathbf{r}''(t)$, es decir, el vector $v'\mathbf{T}$ está comprometida con la variación de la rapidez del movimiento, pero no con la curvatura k de la curva, en tanto que la componente normal de $\mathbf{r}''(t)$, es decir, el vector $kv^2\mathbf{N}$ está comprometida con la curvatura k y con

la rapidez v del movimiento. De aquí vemos que si el movimiento se efectúa sobre una línea recta ($k = 0$), el vector aceleración $\mathbf{r}''(t)$ apuntaría en la dirección del movimiento, pues $\mathbf{r}''(t) = v'\mathbf{T}$, siendo su magnitud igual a

$$\|\mathbf{r}''(t)\| = \|v'\mathbf{T}\| = |v'|\|\mathbf{T}\| = \left|\frac{dv}{dt}\right| = \text{variación de la rapidez del movimiento respecto de } t$$

(esta situación —movimiento rectilíneo— ya era conocida desde los cursos de física elemental). Por otra parte, si el movimiento se realiza de modo que $v =$ cte, entonces el vector aceleración apuntaría en la dirección normal del movimiento, pues en este caso $v' = 0$ y $\mathbf{r}''(t) = kv^2\mathbf{N}$. Esto es justamente lo que ocurre en el llamado "movimiento circular uniforme" (en un plano) estudiado en los cursos de física. En efecto, si $\mathbf{r}(t)$ describe el círculo $x^2 + y^2 = R^2$ con movimiento uniforme (i.e. tal que su velocidad sea constante en magnitud) entonces $v(t) = \|\mathbf{r}'(t)\| =$ cte. y entonces el vector aceleración $a(t) = \mathbf{r}''(t)$ será $a(t) = kv^2\mathbf{N}$. En este caso sabemos que la curvatura k es igual a $\frac{1}{R}$, de modo que

$$a(t) = \frac{v^2}{R}\mathbf{N}$$

donde v es la magnitud de la velocidad del cuerpo moviéndose en el círculo $x^2 + y^2 = R^2$ (la cual es constante, figura 3).

En general (si $k \neq 0$ y $v' \neq 0$), el vector aceleración $\mathbf{r}''(t)$ tendrá sus dos componentes en la dirección de $\mathbf{T}$ y $\mathbf{N}$ no nulas. Reflexionemos un poco más en este hecho. Si vemos a $\mathbf{r}''(t)$ como la rapidez de variación del vector velocidad $\mathbf{r}'(t)$ respecto de t, el cual está cambiando en magnitud (pues $v' \neq 0$) y en dirección (pues $k \neq 0$) vemos que la descomposición de este "vector de variación del vector velocidad" como $v'\mathbf{T}+kv^2\mathbf{N}$, no es más que la separación de la información de la variación de la velocidad *en magnitud* (dada por la componente tangencial $v'\mathbf{T}$) y de la variación de la velocidad *en dirección* (dada por la componente normal $kv^2\mathbf{N}$).

Ejemplo 1. Consideremos el movimiento de un punto en el plano dado por $\mathbf{r}(t) = (t, t^2)$. El punto se mueve entonces por la parábola $y = x^2$ con una velocidad $\mathbf{r}'(t) = (1, 2t)$ y aceleración

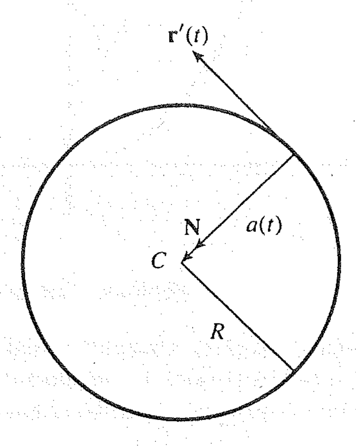


Figura 3. Movimiento circular uniforme.

$\mathbf{r}''(t) = (0, 2)$. Entonces el vector aceleración es constante durante todo el movimiento del punto. Veamos cómo se descompone este vector en sus componentes tangencial y normal. Tenemos que

$$v(t) = \|\mathbf{r}'(t)\| = \sqrt{1 + 4t^2}$$
$$v'(t) = \frac{4t}{\sqrt{1 + 4t^2}}, \qquad k(t) = \frac{2}{(1 + 4t^2)^{3/2}}$$

de modo que

$$\begin{aligned}
\mathbf{r}''(t) &= v'(t)\mathbf{T}(t) + k(t)(v(t))^2\mathbf{N}(t) \\
&= \frac{4t}{\sqrt{1 + 4t^2}}\mathbf{T}(t) + \frac{2}{(1 + 4t^2)^{3/2}}\left(\sqrt{1 + 4t^2}\right)^2 \mathbf{N}(t) \\
&= \frac{4t}{\sqrt{1 + 4t^2}}\mathbf{T}(t) + \frac{2}{\sqrt{1 + 4t^2}}\mathbf{N}(t)
\end{aligned}$$

Se tiene entonces la descomposición

$$(0, 2) = \frac{4t}{\sqrt{1 + 4t^2}}\mathbf{T}(t) + \frac{2}{\sqrt{1 + 4t^2}}\mathbf{N}(t)$$

(la cual se puede verificar poniendo explícitamente a $\mathbf{T}(t)$ y $\mathbf{N}(t)$ que son $\mathbf{T}(t) = \frac{1}{\sqrt{1+4t^2}}(1, 2t)$, $\mathbf{N}(t) = \frac{1}{\sqrt{1+4t^2}}(-2t, 1)$).

Se observa, por ejemplo, que al pasar el punto por el vértice de la parábola ($t = 0$), la componente tangencial de la aceleración se anula (en ese instante la velocidad no cambia en magnitud, sólo en dirección) mientras la componente normal alcanza en ese instante su máximo valor (en magnitud).

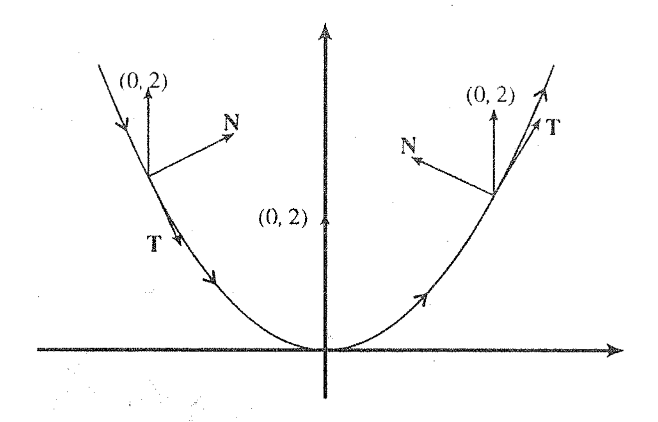


Figura 4. Ejemplo 1. ■

Ejemplo 2. (Movimiento circular). Un punto se mueve en el plano por el círculo $x^2 + y^2 = R^2$ en sentido antihorario, si $\theta(t)$ es el valor del ángulo en radianes medido del vector de posición $\mathbf{r}(t)$ con la parte positiva del eje x, podemos describir el movimiento del punto como

$$\mathbf{r}(t) = (R\cos\theta(t), R\,\mathrm{sen}\,\theta(t))$$

Se define la velocidad angular del punto **p** denotada por $\omega(t)$, como $\omega(t) = \theta'(t)$ = rapidez de variación el ángulo $\theta(t)$ respecto del tiempo t.

Tenemos

$$\begin{aligned}\mathbf{r}'(t) &= (-R\theta'(t)\operatorname{sen}\theta(t), R\theta'(t)\cos\theta(t))\\ &= R\omega(t)(-\operatorname{sen}\theta(t), \cos\theta(t))\end{aligned}$$

de donde

$$v(t) = \|\mathbf{r}'(t)\| = R\omega(t)$$

Así, la aceleración $\mathbf{r}''(t)$ se puede escribir como

$$\begin{aligned}\mathbf{r}''(t) &= v'(t)\mathbf{T}(t) + k(t)(v(t))^2\mathbf{N}(t)\\ &= R\omega'(t)\mathbf{T}(t) + \left(\frac{1}{R}\right)(R^2\omega^2(t))\mathbf{N}(t)\\ &= R\omega'(t)\mathbf{T}(t) + R\omega^2(t)\mathbf{N}(t)\end{aligned}$$

A las componentes tangencial y normal de $\mathbf{r}''(t)$ se les llama *aceleración tangencial* (a_T) y *aceleración normal* (a_N) del movimiento del punto, respectivamente. Así

$$a_T = R\omega'(t)\mathbf{T}(t) \qquad a_N = R\omega^2(t)\mathbf{N}(t)$$

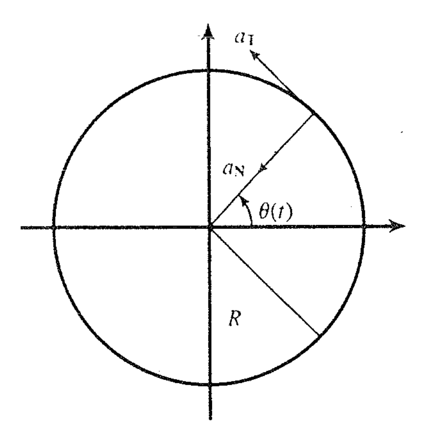


Figura 5. Los vectores a_N y a_T en el movimiento circular

A $\omega'(t)$ se le llama *aceleración angular*, la cual, multiplicada por R nos da la magnitud de la aceleración tangencial.

Como ya habíamos comentado, si el movimiento es uniforme, es decir, $\omega(t) = \omega$ = cte, entonces $\omega'(t) = 0$ y por lo tanto la aceleración tangencial a_T es igual a $R\omega'(t) = R0 = 0$. Así, en este caso, la aceleración del movimiento lo es en la dirección normal y vale $R\omega^2$. Esta nos mide la variación de la dirección del vector $\mathbf{r}'(t)$ —recuerde que su magnitud es $v = R\omega$ = constante. ■

Cuando se estudian movimientos de puntos en el plano, resulta en muchas ocasiones conveniente hacerlo por medio del sistema coordenado polar, cuyo breve estudio iniciamos ahora.

En el sistema cartesiano los puntos se localizan por medio de las coordenadas (cartesianas) x, y; en el sistema polar usamos las coordenadas (polares) ρ y θ donde ρ = distancia del punto al origen, y θ = ángulo que forma el vector posición r del punto con la parte positiva del eje x; este ángulo se determina a menos de múltiplos de 2π. Así la pareja (ρ, θ) determina un único punto del plano, cuyas coordenadas cartesianas (x, y) se relacionan con ρ y θ según las fórmulas

$$x = \rho\cos\theta \qquad y = \rho\,\mathrm{sen}\,\theta$$

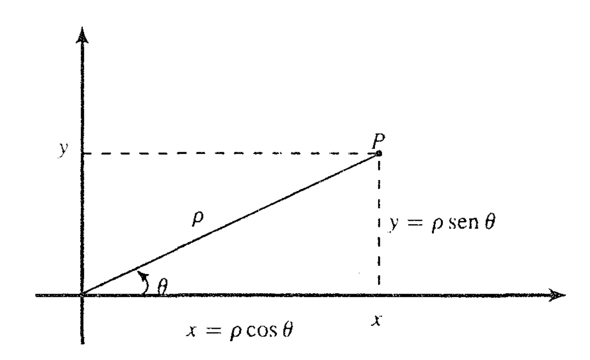


Figura 6. Un punto en el plano polar

Al considerar un camino en el plano $\mathbf{r}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\mathbf{r}(t) = (x(t), y(t))$, podemos escribir las funciones coordenadas $x, y: I \subseteq \mathbb{R} \to \mathbb{R}$ según las fórmulas anteriores como

$$x(t) = \rho(t)\cos\theta(t) \qquad y(t) = \rho(t)\,\mathrm{sen}\,\theta(t)$$

donde ahora $\rho = \rho(t)$ y $\theta = \theta(t)$ serán funciones que a cada valor de $t \in I$ asignan los valores ρ y θ, coordenadas polares del punto $\mathbf{r}(t) \in \mathbb{R}^2$. Escribiremos entonces

$$\mathbf{r}(t) = (\rho(t)\cos\theta(t), \rho(t)\,\mathrm{sen}\,\theta(t)) = \rho(t)(\cos\theta(t), \mathrm{sen}\,\theta(t))$$

Nótese que el vector

$$\mathbf{u}_r(t) = (\cos\theta(t), \mathrm{sen}\,\theta(t))$$

es un vector unitario en la dirección de $\mathbf{r}(t)$. En términos de este vector podemos escribir el camino $\mathbf{r}(t)$ como

$$\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$$

Se introduce también el vector $\mathbf{u}_\theta(t)$ que proviene de girar a $\mathbf{u}_r(t)$ en un ángulo de 90° en dirección antihoraria. Este vector es

$$\mathbf{u}_\theta(t) = (-\,\mathrm{sen}\,\theta(t), \cos\theta(t))$$

Los vectores $\mathbf{u}_r(t)$ y $\mathbf{u}_\theta(t)$ forman el sistema (ortonormal) de referencia respecto del cual conviene expresar la posición $\mathbf{r}(t)$, la velocidad $\mathbf{r}'(t)$ y la aceleración $\mathbf{r}''(t)$, del movimiento descrito por el camino $\mathbf{r}$. (Estos vectores juegan, en el sistema de coordenadas polares, un papel análogo al que los vectores $\mathbf{i} = (1, 0)$ y $\mathbf{j} = (0, 1)$ juegan en el sistema cartesiano: ver el ejemplo 6 de la sección 4, capítulo 1). Hagámoslo.

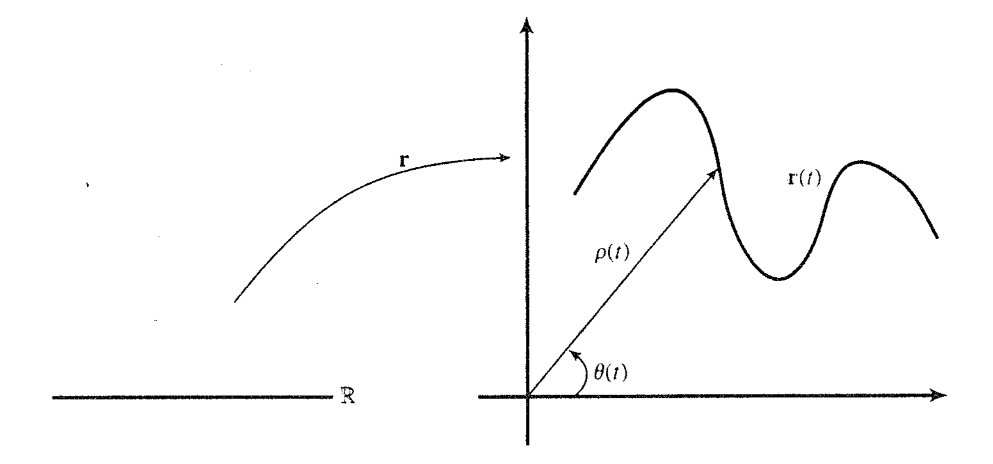


Figura 7. Las funciones $\rho(t)$ y $\theta(t)$.

La posición del punto, ya vimos que se expresa como

$$\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$$

Derivando esta expresión respecto de t obtenemos

$$\mathbf{r}'(t) = \rho(t)\mathbf{u}_r'(t) + \rho'(t)\mathbf{u}_r(t)$$

Nótese que

$$\begin{aligned}\mathbf{u}_r'(t) &= \frac{d}{dt}\big(\cos\theta(t), \operatorname{sen}\theta(t)\big) = \big(-\theta'(t)\operatorname{sen}\theta(t), \theta'(t)\cos\theta(t)\big)\\ &= \theta'(t)\big(-\operatorname{sen}\theta(t), \cos\theta(t)\big) = \theta'(t)\mathbf{u}_\theta(t)\end{aligned}$$

de modo que

$$\mathbf{r}'(t) = \rho'(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)\mathbf{u}_\theta(t)$$

A los factores $\rho'(t)$ y $\rho(t)\theta'(t)$ que aparecen multiplicando a los vectores $\mathbf{u}_r(t)$ y $\mathbf{u}_\theta(t)$ en esta expresión de descomposición ortogonal del vector velocidad $\mathbf{r}'(t)$, se les llama componentes *radial* y *transversal* del vector velocidad $\mathbf{r}'(t)$

La rapidez del movimiento es $v = \|\mathbf{r}'(t)\|$, y se obtiene de la fórmula anterior para $\mathbf{r}'(t)$ recordando que $\mathbf{u}_r$ y $\mathbf{u}_\theta$ son vectores ortonormales

$$\begin{aligned}v = \|\mathbf{r}'(t)\| &= [\,\mathbf{r}'(t)\cdot\mathbf{r}'(t)\,]^{1/2}\\ &= \big[(\rho'(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)\mathbf{u}_\theta(t))\cdot(\rho'(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)\mathbf{u}_\theta(t))\big]^{1/2}\\ &= \sqrt{(\rho'(t))^2 + (\rho(t)\theta'(t))^2}\end{aligned}$$

Antes de obtener la expresión del vector aceleración $\mathbf{r}''(t)$ descompuesto en términos del sistema ortonormal $\mathbf{u}_r$ y $\mathbf{u}_\theta$, veamos un par de ejemplos.

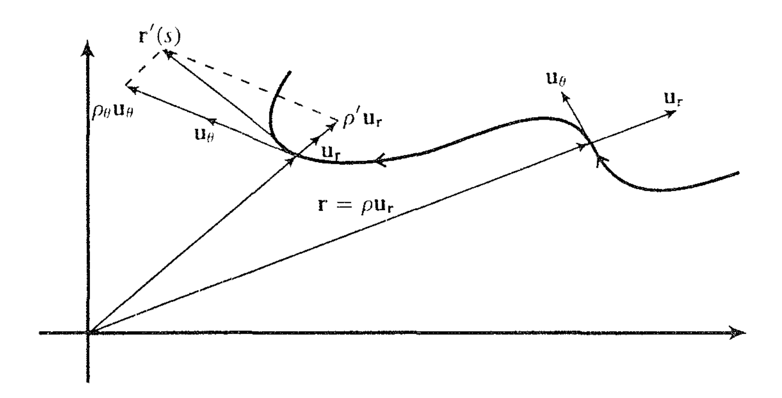


Figura 8. La descomposición del vector $\mathbf{r}'(t)$

Ejemplo 3. Es usual encontrar descripciones de curvas en el sistema polar por medio de una fórmula explícita del tipo $\rho = f(\theta)$. En este caso, poniendo $\theta = \theta(t) = t$, nos queda $\rho(t) = f(t)$, y así, con estas funciones podemos usar las fórmulas de la discusión previa para calcular la velocidad del camino $\mathbf{r}(t) = \rho(t)\mathbf{u}_r = f(t)\mathbf{u}_r$. En este caso particular se obtienen algunas simplificaciones del hecho de que $\theta'(t) = 1$, quedándonos

$$\mathbf{r}'(t) = f'(t)\mathbf{u}_r(t) + f(t)\mathbf{u}_\theta(t)$$

y

$$v(t) = \|\mathbf{r}'(t)\| = \sqrt{(f'(t))^2 + (f(t))^2}$$

o bien, como $\rho = f(\theta) = f(t)$, podemos escribir esta última expresión como

$$v = \sqrt{(f'(\theta))^2 + (f(\theta))^2}$$

De estas observaciones se deduce, mediante la definición de longitud de arco para el camino $\mathbf{r}(t)$ estudiada en la sección 5, que si $a \leq \theta \leq b \leq a + 2\pi$, entonces la longitud de arco entre $\theta = a$ y $\theta = b$ para la curva $\rho = f(\theta)$ es

$$s = \int_a^b \sqrt{(f(\theta))^2 + (f'(\theta))^2} d\theta$$

Dejamos como ejercicio para el lector llenar todos los detalles que concluyen la validez de esta fórmula.

La ecuación más simple con que podemos ejemplificar la situación anterior es $\rho = f(\theta) = R$, que corresponde a un círculo con centro en el origen y radio R. En este caso el vector velocidad es

$$\mathbf{r}'(t) = f'(t)\mathbf{u}_r + f(t)\mathbf{u}_\theta = R\mathbf{u}_\theta$$

como ya sabíamos. También la longitud del círculo es, según la fórmula obtenida arriba

$$s = \int_0^{2\pi} \sqrt{(R)^2 + (0)^2} d\theta = 2\pi R$$

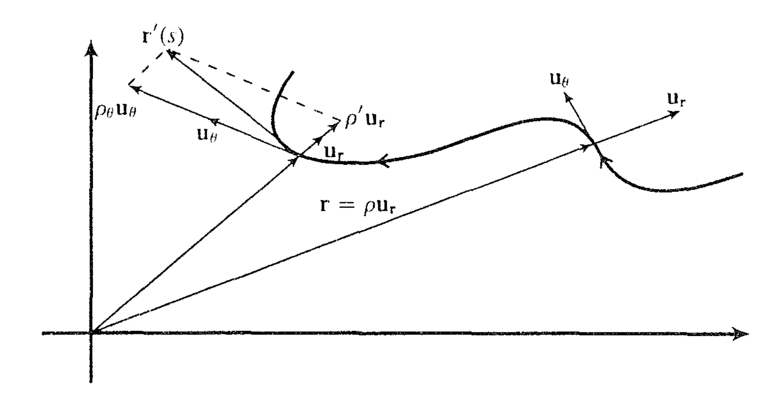


Figura 8. La descomposición del vector $\mathbf{r}'(t)$.

Ejemplo 3. Es usual encontrar descripciones de curvas en el sistema polar por medio de una fórmula explícita del tipo $\rho = f(\theta)$. En este caso, poniendo $\theta = \theta(t) = t$, nos queda $\rho(t) = f(t)$, y así, con estas funciones podemos usar las fórmulas de la discusión previa para calcular la velocidad del camino $\mathbf{r}(t) = \rho(t)\mathbf{u}_r = f(t)\mathbf{u}_r$. En este caso particular se obtienen algunas simplificaciones del hecho de que $\theta'(t) = 1$, quedándonos

$$\mathbf{r}'(t) = f'(t)\mathbf{u}_r(t) + f(t)\mathbf{u}_\theta(t)$$

y

$$v(t) = \|\mathbf{r}'(t)\| = \sqrt{(f'(t))^2 + (f(t))^2}$$

o bien, como $\rho = f(\theta) = f(t)$, podemos escribir esta última expresión como

$$v = \sqrt{(f'(\theta))^2 + (f(\theta))^2}$$

De estas observaciones se deduce, mediante la definición de longitud de arco para el camino $\mathbf{r}(t)$ estudiada en la sección 5, que si $a \leq \theta \leq b \leq a + 2\pi$, entonces la longitud de arco entre $\theta = a$ y $\theta = b$ para la curva $\rho = f(\theta)$ es

$$s = \int_a^b \sqrt{(f(\theta))^2 + (f'(\theta))^2} d\theta$$

Dejamos como ejercicio para el lector llenar todos los detalles que concluyen la validez de esta fórmula.

La ecuación más simple con que podemos ejemplificar la situación anterior es $\rho = f(\theta) = R$, que corresponde a un círculo con centro en el origen y radio R. En este caso el vector velocidad es

$$\mathbf{r}'(t) = f'(t)\mathbf{u}_r + f(t)\mathbf{u}_\theta = R\mathbf{u}_\theta$$

como ya sabíamos. También la longitud del círculo es, según la fórmula obtenida arriba

$$s = \int_0^{2\pi} \sqrt{(R)^2 + (0)^2} d\theta = 2\pi R$$

como también ya sabíamos desde hace mucho tiempo.

Otro ejemplo es la ecuación $\rho = f(\theta) = 1 + \cos\theta$. La curva correspondiente a esta ecuación se llama *cardioide*. Si un punto se mueve sobre ella según la ecuación anterior lo haría con una velocidad igual a

$$\mathbf{r}'(t) = f'(t)\mathbf{u}_r + f(t)\mathbf{u}_\theta = (-\operatorname{sen}\theta)\mathbf{u}_r + (1 + \cos\theta)\mathbf{u}_\theta$$

La longitud total de la cardioide es

$$\begin{aligned} s &= \int_0^{2\pi} \sqrt{(f(\theta))^2 + (f'(\theta))^2}d\theta = \int_0^{2\pi} \sqrt{(1 + \cos\theta)^2 + (-\operatorname{sen}\theta)^2}d\theta \\ &= 2\int_0^{2\pi} \sqrt{2 + 2\cos\theta}d\theta = \int_0^{2\pi} \sqrt{\cos^2(\theta/2)}d\theta = 8 \end{aligned}$$

(ver la integral en el ejemplo 2 de la sección 5)

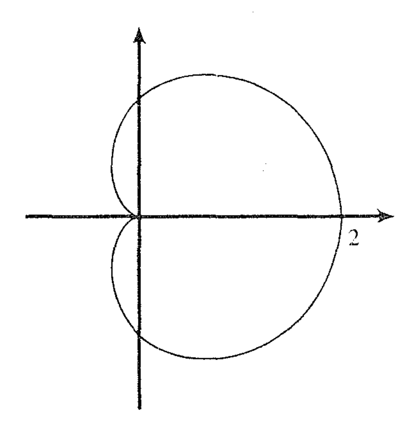


Figura 9. Cardioide. ■

Ejemplo 4. (El problema de las cuatro moscas). Este es un divertido e interesante problema que se refiere a cuatro moscas que un buen día se posaron sobre las cuatro esquinas de una mesa cuadrada (una en cada esquina) de modo que cada una sólo veía a su compañera de la esquina vecina (siguiendo el recorrido de visión de las moscas en sentido antihorario). En un momento dado las cuatro moscas se pusieron a caminar sobre la mesa con la misma velocidad (rapidez) de modo que cada una de ellas siempre veía frente a sí sólo a la compañera que estaba viendo desde el inicio. Se trata de descubrir la trayectoria que seguirán las moscas y estudiar si habrá un momento en que ocurra un feliz encuentro entre ellas, así como la distancia que tendrán que caminar antes de tal encuentro.

Consideremos un sistema coordenado polar con el origen en el centro de la mesa como se muestra en la figura 10.

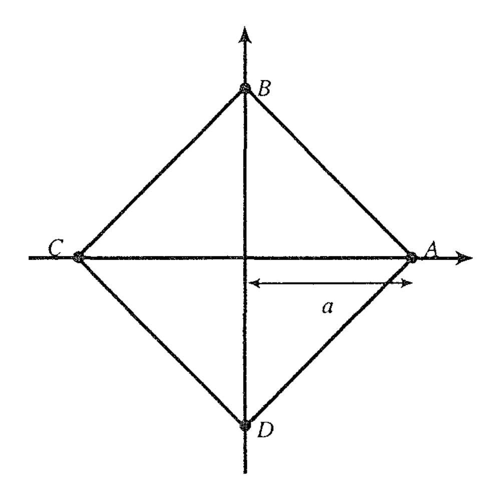


Figura 10. La posición inicial de cada mosca

Resolvamos el problema para la mosca m situada en el punto A (la situación es totalmente simétrica para las tres cuatro moscas). Al tiempo t la mosca m se encontrará en $\mathbf{r}(t)$ donde

$$\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t) = \rho(t)(\cos\theta(t), \operatorname{sen}\theta(t))$$

en tanto que la mosca que se encontraba en B se *debe* encontrar en el punto

$$\rho(t)\left(\cos\left(\frac{\pi}{2} + \theta(t)\right), \operatorname{sen}\left(\frac{\pi}{2} + \theta(t)\right)\right) = \rho(t)(-\operatorname{sen}\theta(t), \cos\theta(t)) = \rho(t)\mathbf{u}_\theta(t)$$

El vector velocidad $\mathbf{r}'(t)$ debe obedecer entonces que

$$\mathbf{r}'(t) = \rho(t)\mathbf{u}_\theta(t) - \rho(t)\mathbf{u}_r(t)$$

(esta es la condición que nos dice que la mosca m se está moviendo de modo que siempre está viendo a la compañera que estaba situada en B).

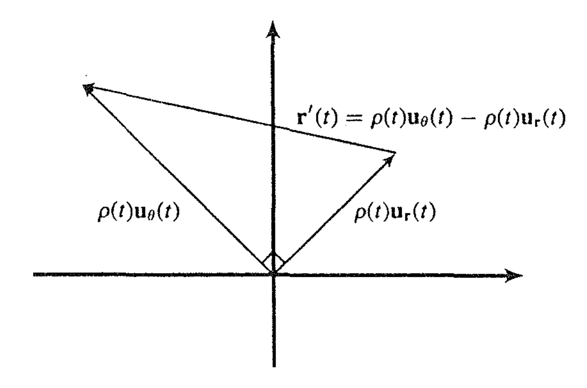


Figura 11. El movimiento de una de las moscas.

Se tiene entonces

$$\mathbf{r}'(t) = \rho'(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)\mathbf{u}_\theta(t) = \rho(t)\mathbf{u}_\theta(t) - \rho(t)\mathbf{u}_r(t)$$

o sea

$$(\rho'(t) + \rho(t))\mathbf{u}_r(t) + \rho(t)(\theta'(t) - 1)\mathbf{u}_\theta(t) = 0$$

Esta es una relación que se debe cumplir para todo $t > 0$. Entonces

$$\rho'(t) + \rho(t) = 0$$
$$\rho(t)(\theta'(t) - 1) = 0$$

(¿por qué?). De la segunda expresión se obtiene que (puesto que $\rho(t) \neq 0$) $\theta'(t) = 1$, o sea $\theta(t) = t = k_1$, donde k_1 es una constante. Con $\theta(0) = 0$, se tiene $k_1 = 0$. Así que $\theta(t) = t$. De la primera expresión se obtiene

$$\frac{d\rho}{dt} + \rho = 0$$

o bien

$$\frac{d\rho}{\rho} = -dt = -d\theta$$

Integrando se obtiene

$$\ln \rho = -\theta + k_2$$

como en $t = \theta = 0$ se tiene $\rho = a$, vemos que $k_2 = \ln a$, de modo que, sustituyendo este valor en la expresión anterior y despejando ρ en términos de θ llegamos a

$$\rho = ae^{-\theta}$$

Esta es la ecuación que describe la trayectoria de la mosca m. Se trata de una *espiral logarítmica*. Cuando $t = \theta \to \infty$, se tiene $\rho \to 0$. Entonces en el origen es donde ocurrirá el encuentro de las cuatro moscas. Teóricamente transcurrirá un tiempo infinito para que ocurra este encuentro (solamente en $t = \theta = \infty$ se tiene $\rho = 0$). Sin embargo, vemos que la distancia que las moscas caminarán hasta su encuentro es finita. Esta será la longitud de la curva $\rho = f(\theta) = ae^{-\theta}$ desde $\theta = 0$ hasta $\theta = \infty$, es decir

$$s = \int_0^\infty \sqrt{(\mathbf{f}(\theta))^2 + (\mathbf{f}'(\theta))^2}d\theta = \int_0^\infty \sqrt{(ae^{-\theta})^2 + (-ae^{-\theta})^2}d\theta$$
$$= \int_0^\infty \sqrt{2a^2e^{-2\theta}}d\theta = \sqrt{2}a \int_0^\infty e^{-\theta}d\theta = -\sqrt{2}ae^{-\theta}\Big|_0^\infty = \sqrt{2}a$$

Así, cada una de las moscas caminará sobre una espiral logarítmica del tipo $\rho = ae^{-\theta}$ (esta es la ecuación que le corresponde a la mosca que arrancó del punto A) y después de un "tiempo infinito" las moscas se encontrarán en el centro de la mesa, habiendo recorrido cada una de ellas una distancia de $\sqrt{2}a$=longitud del lado de la mesa.

Pasemos ahora a estudiar la aceleración de un punto moviéndose en el camino $\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$. Habíamos obtenido que

$$\mathbf{r}'(t) = \rho'(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)\mathbf{u}_\theta(t)$$

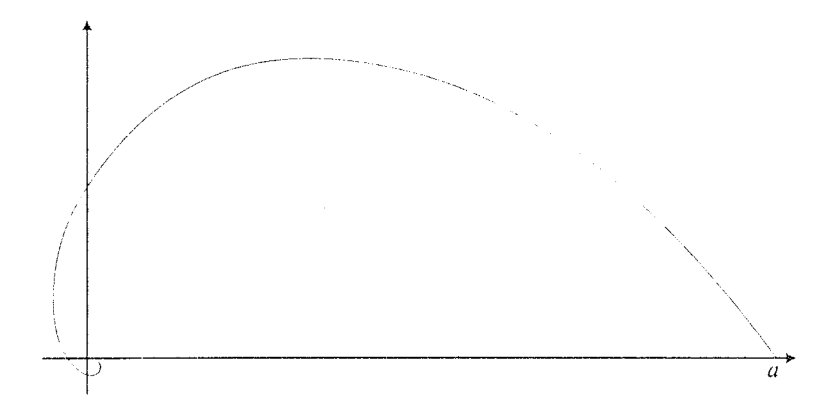


Figura 12. La espiral logarítmica $\rho = ae^{-\theta}$ ■

Una nueva derivación nos conduce a

$$\mathbf{r}''(t) = \rho'(t)\mathbf{u}_r'(t) + \rho''(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)\mathbf{u}_\theta'(t) + \rho(t)\theta''(t)\mathbf{u}_\theta(t) + \rho'(t)\theta'(t)\mathbf{u}_\theta(t)$$

Sabemos que $\mathbf{u}_r'(t) = \theta'(t)\mathbf{u}_\theta(t)$. Un cálculo directo muestra también que $\mathbf{u}_\theta'(t) = -\theta'(t)\mathbf{u}_r(t)$, de modo que la expresión anterior queda como

$$\mathbf{r}''(t) = \rho'(t)\theta'(t)\mathbf{u}_\theta(t) + \rho''(t)\mathbf{u}_r(t) + \rho(t)\theta'(t)(-\theta'(t)\mathbf{u}_r(t)) + \rho(t)\theta''(t)\mathbf{u}_\theta(t) + \rho'(t)\theta'(t)\mathbf{u}_\theta(t)$$

o bien

$$\mathbf{r}''(t) = (\rho''(t) - \rho(t)(\theta'(t))^2\mathbf{u}_r(t) + (\rho(t)\theta''(t) + 2\rho'(t)\theta'(t))\mathbf{u}_\theta(t)$$

que es la expresión que nos da la aceleración $\mathbf{r}''(t)$ descompuesta en la suma de sus componentes $a_r(t) = \rho''(t) - \rho(t)(\theta'(t))^2$ en la dirección de $\mathbf{u}_r(t)$, llamada componente radial, y $a_\theta(t) = \rho(t)\theta''(t) + 2\rho'(t)\theta'(t)$ en la dirección de $u_\theta(t)$, llamada componente transversal.

Ejemplo 5. Si en la fórmula anterior para $\mathbf{r}''(t)$ ponemos $\theta(t) = t$, ésta se ve como

$$\mathbf{r}''(t) = (\rho''(t) - \rho(t))\mathbf{u}_r(t) + 2\rho'(t)\mathbf{u}_\theta(t)$$

Por ejemplo, si un punto se mueve sobre la cardioide $\rho = \mathbf{f}(\theta) = 1 + \cos\theta$, poniendo $\theta = \theta(t) = t$, nos queda $\rho = \rho(t) = 1 + \cos t$. En el ejemplo 3 anterior, vimos que la velocidad del punto es

$$\mathbf{r}'(t) = (-\operatorname{sen} t)\mathbf{u}_r + (1 + \cos t)\mathbf{u}_\theta$$

La aceleración $\mathbf{r}''(t)$ del movimiento de este punto será entonces

$$\begin{aligned}\mathbf{r}''(t) &= (\rho''(t) - \rho(t))\mathbf{u}_r(t) + 2\rho'(t)\mathbf{u}_\theta(t)\\ &= (-\cos t - 1 - \cos t)\mathbf{u}_r(t) + 2(-\operatorname{sen} t)\mathbf{u}_\theta(t)\\ &= -(1 + 2\cos t)\mathbf{u}_r(t) - 2\operatorname{sen} t\mathbf{u}_\theta(t)\end{aligned}$$

La rapidez con que se mueve el punto es

$$v = \|\mathbf{r}'(t)\| = \sqrt{(-\operatorname{sen} t)^2 + (1 + \cos t)^2} = \sqrt{2 + 2\cos t}$$

y la magnitud de la aceleración del movimiento es

$$a = \|\mathbf{r}''(t)\| = \sqrt{(1 + 2\cos t)^2 + (2\,\mathrm{sen}\, t)^2} = \sqrt{5 + 4\cos t}$$ ■

Un caso particular muy interesante de movimiento de un punto en el plano se presenta cuando la componente transversal de la aceleración es igual a cero. En tal caso decimos que la aceleración del movimiento es una *aceleración radial*. La propiedad más importante que se presenta en el movimiento de un punto con aceleración radial se recoge en el siguiente teorema.

Teorema 5.11.1 Supongamos que un punto se mueve en el plano según el camino $\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$, con aceleración radial. Sea $A(t)$ el área barrida por el vector $\mathbf{r}(t)$ entre el instante $t = t_0$ (dado) y el instante $t > t_0$. Entonces $A(t)$ es proporcional al tiempo t.

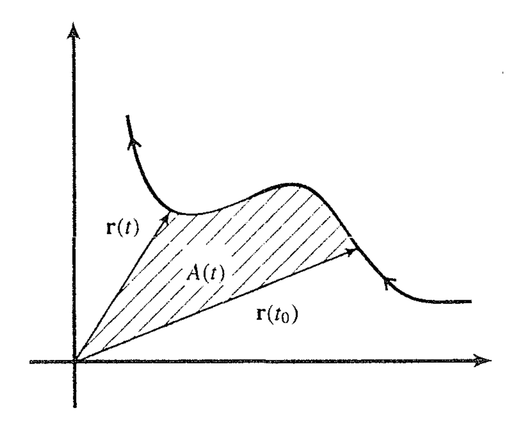


Figura 13. El área que barre el vector $\mathbf{r}(t)$ con aceleración radial. ■

Demostración. Siendo la aceleración radial se tiene

$$a_\theta(t) = \rho(t)\theta''(t) + 2\rho'(t)\theta'(t) = 0$$

es decir

$$\frac{1}{\rho(t)}\frac{d}{dt}\left((\rho(t))^2\theta'(t)\right) = 0$$

de donde se deduce que

$$(\rho(t))^2\theta'(t) = m = \text{cte.} \qquad (*)$$

Por otra parte, sea $\rho = f(\theta)$ la expresión explícita de la curva descrita por el camino $\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$. Entonces $\rho(t) = f(\theta(t))$. Sabemos que el área A de la curva $\rho = f(\theta)$ entre θ_0 y θ_1 se calcula como

$$A = \frac{1}{2}\int_{\theta_0}^{\theta_1} (f(\theta))^2 d\theta$$

En nuestro caso $\theta_0 = \theta(t_0)$ y $\theta_1 = \theta(t)$, así que A es una función de θ_1 y por lo tanto de t. Tenemos

$$\begin{aligned} A'(t) &= \frac{dA}{dt} = \frac{dA}{d\theta_1}\frac{d\theta_1}{dt} = \left(\frac{d}{d\theta_1}\left(\frac{1}{2}\int_{\theta_0}^{\theta_1}(f(\theta))^2\,d\theta\right)\right)\theta'(t) \\ &= \left(\frac{1}{2}f(\theta_1)^2\right)\theta'(t) = \left(\frac{1}{2}f(\theta(t))^2\theta'(t)\right) = \frac{1}{2}(\rho(t))^2\theta'(t) \end{aligned}$$

Entonces, según (*) tenemos que $A'(t) = \bar{m}$, ($\bar{m} = 2m$) de donde $A(t) = \bar{m}t + C$, es decir, $A(t)$ es proporcional a t. Q.E.D.

Terminamos esta sección con algunos comentarios sobre la descripción del movimiento de los planetas.

La llamada "Ley de la Gravitación Universal", debida a Sir Isaac Newton (1642–1727), establece que la fuerza de atracción entre dos cuerpos en el espacio de masas m_1 y m_2 es directamente proporcional al producto de sus masas e inversamente proporcional al cuadrado de la distancia que las separa (medida ésta desde los centros de masa de los cuerpos). Digamos que esta distancia es d. Entonces la fuerza F viene dada por

$$F = G\frac{m_1 m_2}{d^2}$$

donde G es la constante de proporcionalidad de la ley, llamada constante de la gravitación universal (que vale aproximadamente 6.67×10^{-11} $\mathrm{Nm^2/kg^2}$).

Consideremos el movimiento de un planeta alrededor del sol. Sea M la masa del sol y m la masa del planeta. Sea también $\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$ el camino (en $\mathbb{R}^3$) que describe la trayectoria del planeta, suponiendo que el sol se encuentra en el origen de cordenadas. La segunda ley de Newton establece que la fuerza con que el planeta es atraído hacia el sol es $F = ma$, donde a es la aceleración del planeta. Combinando esta expresión con la correspondiente a la Ley de Gravitación Universal, que en este caso es $F = G\frac{mM}{d^2}$, donde $d = \|\mathbf{r}(t)\|$ = distancia del planeta al sol, nos queda

$$a = \frac{GM}{d^2}$$

Vista en forma vectorial, esta aceleración está apuntando al origen de coordenadas. Es decir, su dirección es la del vector $-\mathbf{u}_r$. Podemos expresar entonces la aceleración $\mathbf{a}(t)$ vectorialmente como

$$\mathbf{a}(t) = -\frac{GM}{d^2}\mathbf{u}_r(t)$$

donde $d = \|\mathbf{r}(t)\|$

Estas ideas sugieren inmediatamente que la aceleración del planeta es una "aceleración radial", como la consideramos en el teorema 5.11.1. Recuerde que tal concepto lo establecimos para el movimiento de un punto en el plano, en tanto que nuestro planeta se está moviendo en el espacio. Sin embargo, el siguiente resultado nos dice que, de hecho, el movimiento del planeta se realiza en un plano.

Lema Si el movimiento de un punto en el espacio descrito por el camino $\mathbf{r}(t) = \rho(t)\mathbf{u}_r(t)$ es tal que su aceleración $\mathbf{a} = \mathbf{r}''(t)$ es

$$\mathbf{a} = -\frac{GM}{d^2}\mathbf{u}_r(t), \quad d = \|\mathbf{r}(t)\|$$

entonces el punto se mueve sobre un plano (que pasa por el origen). ■

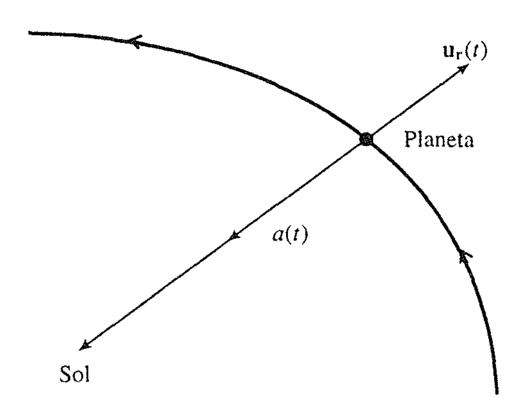


Figura 14. El movimiento de un planeta alrededor del Sol.

Demostración. Se tiene

$$\begin{aligned}\frac{d}{dt}(\mathbf{r}(t)\times\mathbf{r}'(t)) &= \mathbf{r}(t)\times\mathbf{r}''(t)+\mathbf{r}'(t)\times\mathbf{r}'(t)=\mathbf{r}(t)\times\mathbf{r}''(t)+0\\ &= (\rho(t)\mathbf{u}_r(t))\times\left(-\frac{GM}{d^2}\mathbf{u}_r(t)\right)\\ &= -\frac{GM\rho(t)}{d^2}\mathbf{u}_r(t)\times\mathbf{u}_r(t)=0\end{aligned}$$

Entonces $\mathbf{r}(t)\times\mathbf{r}'(t)=\gamma$, en que γ es un vector constante (no depende de t). Si $\gamma=0$, entonces $\mathbf{r}(t)$ y $\mathbf{r}'(t)$ serían vectores paralelos, lo cual nos indicaría que el movimiento del punto es rectilíneo y por lo tanto es plano. Si $\gamma\neq 0$, entonces tomando producto punto con el vector $r(t)$ en ambos miembros de la expresión $\mathbf{r}(t)\times\mathbf{r}'(t)=\gamma$, nos queda $\gamma\cdot\mathbf{r}(t)=(\mathbf{r}(t)\times\mathbf{r}'(t))\cdot\mathbf{r}(t)=0$. Entonces $\mathbf{r}(t)$ es un vector (vector posición del punto) que siempre es ortogonal a γ; es decir, $\mathbf{r}(t)$ se encuentra siempre en un plano en $\mathbb{R}^3$ con vector normal γ. Q.E.D.

Con este lema podemos considerar entonces el movimiento de un planeta alrededor del sol, (según la ley de la gravitación universal y la segunda ley de Newton) como el movimiento de un punto en el plano (el plano obtenido en el lema, con el sol en el origen de coordenadas), *con aceleración radial*. Por lo tanto, a la luz del teorema 5.11.1, podemos concluir que el movimiento del planeta alrededor del sol se realiza de manera que barre áreas iguales en tiempos iguales.
En la figura anterior, si $t_2-t_1=t_4-t_3$ entonces $A_1=A_2$. Este es un resultado célebre obtenido por el astrónomo alemán Johannes Kepler (1571–1630) quien solamente contaba con los datos astronómicos de que se disponía en su tiempo, conocido como "segunda ley de Kepler". Hay otras leyes que Kepler obtuvo también en forma empírica. Estas son:

1a. ley de Kepler: los planetas desciben órbitas elípticas en uno de cuyos focos está el sol.

3a. ley de Kepler: el cuadrado del periodo de un planeta (tiempo que tarda en dar una vuelta completa al sol) es proporcional al cuadrado de su distancia media al sol.

Estas leyes también se pueden obtener "matemáticamente", como lo hicimos con la segunda ley. Es decir, estas leyes resultan ser una consecuencia de la Ley de Gravitación Universal de Newton (ver, por ejemplo, [ApI], pp. 669, 670, 671).

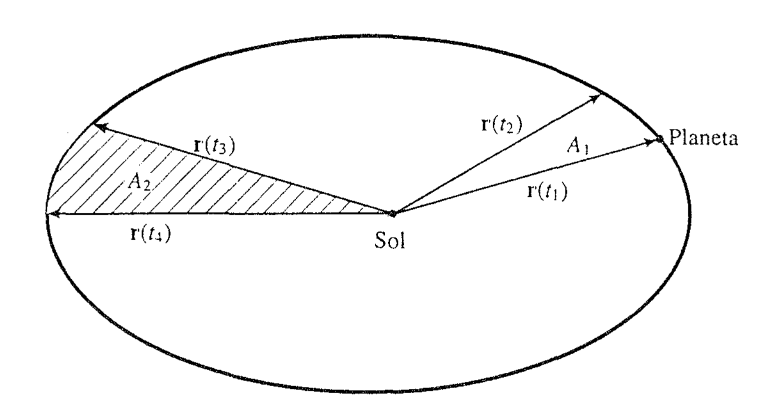


Figura 15. Áreas que un planeta barre en su movimiento alrededor del Sol

Ejercicios (Capítulo 5, Sección 11)

En los ejercicios 1–5, descomponga el vector aceleración de la curva dada en el punto indicado, como la suma de sus componentes tangencial y normal.

1. $\mathbf{r}(t) = (3t + 1, t - 2)$, en un punto cualquiera $\mathbf{p} = \mathbf{r}(t)$.

2. $\mathbf{r}(t) = (t, t^3)$, en el punto $\mathbf{p} = \mathbf{r}(1)$.

3. $\mathbf{r}(t) = (t, t^3)$, en el punto $\mathbf{p} = \mathbf{r}(-1)$.

4. $\mathbf{r}(t) = (2\cos t, \operatorname{sen} t)$, en el punto $\mathbf{p} = \mathbf{r}(0)$.

5. $\mathbf{r}(t) = (t^2, t^3)$, en el punto $\mathbf{p} = \mathbf{r}(1)$.

6. Considere el camino $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = ((1 + \cos t)\cos t, (1 + \cos t)\operatorname{sen} t))$. Determine los vectores velocidad $\mathbf{f}'(t)$ y aceleración $\mathbf{f}''(t)$ en el punto $\mathbf{p} = \mathbf{f}(0)$.

7. Determine los vectores velocidad $\mathbf{r}'(t)$ y aceleración $\mathbf{r}''(t)$ para un punto que se mueve sobre la cardioide $r = 1 + \cos\theta$, en el punto correspondiente a $\theta = 0$ (ver ejemplo 3). Compare con el ejercicio anterior. ¿Cómo son las curvas de estos dos ejercicios?

8. Sea $r = r(\theta)$ la ecuación de una curva en coordenadas polares, con a $\leq \theta \leq b$. Demuestre que su curvatura se puede escribir como

$$k(\theta) = \frac{2(r'(\theta))^2 - r(\theta)r''(\theta) + (r(\theta))^2}{\left((r'(\theta))^2 + (r(\theta))^2\right)^{3/2}}$$

9. Calcule la curvatura de la espiral logarítmica $r = e^{-\theta}$ en un punto cualquiera. ¿Qué sucede cuando θ tiende a infinito?

10. Calcule la curvatura de la espiral de Arquímedes $r = \theta$ en un punto cualquiera. Compare con el ejercicio anterior.

11. Calcule el radio de curvatura de la cardioide $r = 1 + \cos\theta$, en el punto correspondiente a $\theta = 0$.

Capítulo 6

Integrales múltiples

En este capítulo estudiaremos el cálculo integral de las funciones de varias variables. El objetivo del capítulo es estudiar las integrales de funciones del tipo $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ sobre algunos subconjuntos D de U, así como algunas de las aplicaciones de estas integrales en problemas de geometría (cálculo de áreas y volúmenes de cuerpos en el plano y en el espacio) y de mecánica (cálculo de centros de masa de cuerpos en el plano y en el espacio).

A manera de introducción, recordemos que una de las ideas que acompañaron al cálculo integral de funciones de una sola variable $f: I \subseteq \mathbb{R} \to \mathbb{R}$ fue la del "área bajo la curva". En efecto, en el primer curso de cálculo se vió que "en determinadas circunstancias", la integral de la función $y = f(x)$ sobre el intervalo $[a, b]$ se podía interpretar como el cálculo del área de la figura en el plano, limitada por las rectas $x = a$ (por la izquierda), $x = b$ (por la derecha), $y = 0$ (por abajo) y la gráfica de la función $y = f(x)$ (por arriba). Nos referíamos entonces a esta área como "el área bajo la curva de $y = f(x)$ entre a y b". Las "circunstancias" anteriormente mencionadas eran que la función tenía que ser positiva (su gráfica debía quedar por encima del eje x) y continua.

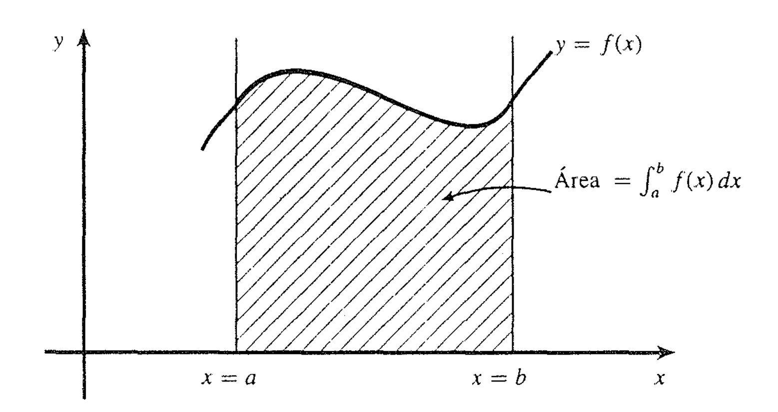


Figura 1. La integral de $y = f(x)$ entre $x = a$ y $x = b$ es el área bajo la curva.

En la generalización que haremos de este concepto para integrales de funciones de dos variables $z = f(x, y)$, integrales que se harán ahora no sobre "partes de la recta" (como el intervalo $[a, b]$) como se hacía para funciones de una sola variable, sino sobre "partes del plano" (una parte del dominio de la función $z = f(x, y)$), la idea análoga que surgirá será la de "volumen bajo la superficie". Más

en concreto, si tenemos una función $z = f(x, y)$ (supongámosla por el momento definida en todo $\mathbb{R}^2$) continua y positiva (que la superficie que representa su gráfica esté por encima del plano xy), entonces la integral que definiremos de esta función sobre un subconjunto $D \subseteq \mathbb{R}^2$ será el volumen del "cilindro" limitado por el plano xy (por abajo), la gráfica de la función $z = f(x, y)$ (por arriba), y la frontera de D marcando la parte lateral del cuerpo resultante. A esta integral la representaremos como

$$\iint_D f(x, y)\, dx\, dy$$

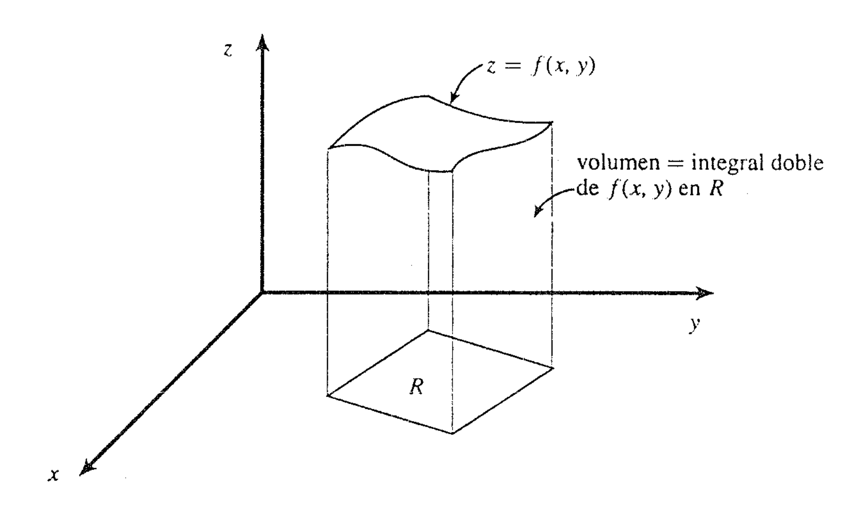


Figura 2. La integral doble de $z = f(x, y)$ como volumen bajo la superficie.

El desarrollo del capítulo se dividirá en tres partes fundamentales. En la primera de ellas se estudiarán de manera exhaustiva las integrales dobles. Se hará al comienzo (secciones 1 y 2) una presentación "matemáticamente decente" del concepto de integral doble, la cual es de hecho la misma que se hace después para las integrales triples y para las n-múltiples en general. Esta presentación es fundamentalmente teórica y se podría evitar en un primer acercamiento a este tema. Como suele suceder en matemáticas, una vez establecido el concepto de integral doble, con todos los detalles teóricos que la buena educación matemática demanda, se verá que aunque muy bien sustentado el concepto, su definición resulta ciertamente inútil para *calcular* tales integrales (como llega a suceder con las integrales de funciones de una sola variable: la definición de la integral como límite de sumas de Riemann resulta impráctica para calcular esas integrales. El Teorema Fundamental del Cálculo nos rescata de tal situación y nos proporciona un modo mucho más agradable de efectuar cálculos de integrales definidas). En la sección 3 se verá que la situación no es tan grave como parece; en la práctica, las integrales dobles no se calculan con la definición dada (¡afortunadamente!). Se verá que las integrales dobles de muchas funciones importantes se podrán calcular como "dos integrales simples" (donde "integral simple" quiere decir "integral de una función de una sola variable" y entonces, con el Teorema Fundamental del Cálculo se podrán hacer estos cálculos de manera sencilla). Más aún, en la sección 4 se presentan los poderosos "trucos" de los cambios de variables para hacer lo más divertido posible (o si el lector así lo quiere ver, lo menos tortuoso posible) el cálculo de

estas integrales. Hacia el final de esta primera parte (en la sección 5) se estudiarán algunas de las aplicaciones de las integrales dobles. En la segunda parte del capítulo se estudian las integrales triples (secciones 6,7 y 8). El esquema con el que se aborda este estudio es completamente análogo al desarrollo de las secciones 1 a 5 correspondientes al caso de las integrales dobles y , en general, las discusiones son mucho más escuetas, pues habiendo entendido qué pasa con las integrales dobles, en algunos casos entender lo que pasa con las integrales triples resulta un proceso trivial: basta escribir un signo de integral más y una letra más en las variables de las funciones involucradas. En la tercera parte del capítulo (la sección 9) se presenta, "en resumidas cuentas", lo que pudo haberse hecho desde el comienzo: el estudio de las integrales n-múltiples en general. En esta sección se vuelve a repetir el esquema del estudio sistemático que se hizo con las integrales dobles en las secciones 1 a 5, sólo que, claro está, de manera mucho más resumida. Esta sección es opcional, puede ser evitada sin peligro alguno de perder la secuencia de temas para los capítulos restantes. Está escrita para quienes disfrutan de las generalizaciones en matemáticas. En ella se presentan los cálculos correspondientes para hallar los "volúmenes" de esferas, conos y paralelepípedos en el espacio $\mathbb{R}^n$.

6.1 Integrales dobles (I): funciones escalonadas

El objetivo de esta sección es establecer el concepto de integral doble para determinado tipo de funciones sencillas llamadas funciones escalonadas. Ciertamente este tipo de funciones no aparecen frecuentemente en la práctica, pero su estudio nos permitirá, por una parte, captar la esencia del concepto de un integral doble y , por otra, basándonos en las integrales de estas funciones se construirá el concepto de integral doble para funciones más generales.

Comencemos por introducir los dominios sobre los que están definidas las funciones escalonadas, que serán también las regiones sobre las cuales integraremos estas funciones.

Un *rectángulo* en $\mathbb{R}^2$, que denotaremos por Q, se define como el producto cartesiano de dos intervalos de $\mathbb{R}$, digamos I_1 e I_2. Es decir

$$Q = I_1 \times I_2 = \{(x, y) \in \mathbb{R}^2 | x \in I_1, y \in I_2\}$$

Los intervalos I_1 e I_2 pueden ser abiertos, cerrados, etc. (y en cada caso se dice que el rectángulo Q correspondiente es abierto, cerrado, etc.). La diferencia entre el rectángulo abierto $Q_a = (a, b) \times (c, d)$ y el rectángulo cerrado $Q_c = [a, b] \times [c, d]$ es, digámosolo así, "la orilla" del rectángulo. Las funciones que vamos a estudiar en esta sección estarán definidas en rectángulos y, para la teoría a desarrollar, el valor de tales funciones en las orillas de los rectángulos en que están definidas es completamente irrelevante. Aquí, sólo para fijar ideas, vamos a considerar siempre rectángulos cerrados.

Si $I_1 = [a, b]$ e $I_2 = [c, d]$, se tiene entonces que el rectángulo $Q = I_1 \times I_2$ es

$$Q = [a, b] \times [c, d] = \{(x, y) \in \mathbb{R}^2 | x \in [a, b], y \in [c, d]\}$$

el cual se ve así en forma geométrica (figura 1).

Recuerde que una partición P de un intervalo $I = [a, b]$ de la recta de los reales es un conjunto de puntos $x_0, x_1, \ldots, x_n$ de I tales que

$$a = x_0 < x_1 < \ldots x_{n-1} < x_n = b$$

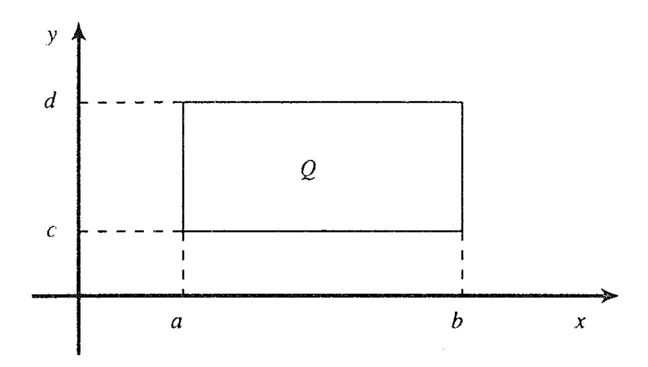


Figura 1. El rectángulo $[a, b] \times [c, d]$ en el plano.

Esquemáticamente tenemos

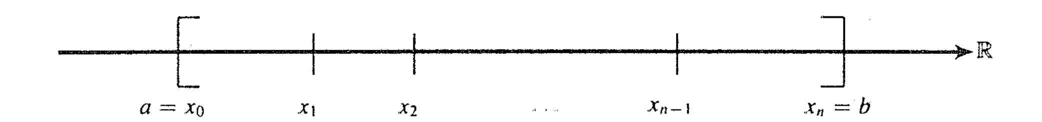


Figura 2. Una partición P del intervalo $[a, b]$.

Una partición P de un rectángulo $Q = I_1 \times I_2$ en $\mathbb{R}^2$ es un producto del tipo $P = P_1 \times P_2$ donde P_j es una partición del intervalo I_j, $j = 1, 2$. Obsérvese entonces que podemos pensar el rectángulo Q como dividido por la partición P en nm subrectángulos, donde n (m, respectivamente) es el número de subintervalos en que quedó divido el intervalo I_1 (I_2, respectivamente) por la partición P_1 (P_2, respectivamente). Más en concreto, suponga que los intervalos que conforman Q son $I_1 = [a, b]$ e $I_2 = [c, d]$ y que las particiones correspondientes P_1 y P_2 son

$$P_1 = \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\}$$
$$P_2 = \{c = y_0 < y_1 < y_2 < \ldots < y_m = d\}$$

Entonces

$$P = P_1 \times P_2 = \{(x_i, y_j) | x_i \in P_1,\ y_j \in P_2,\ i = 0, 1, \ldots, n,\ j = 0, 1, \ldots, m\}$$

y los nm subrectángulos de Q que produce esta partición son

$$Q_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \qquad i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, m$$

Como ejemplo concreto, tomemos el rectángulo Q de la figura 1 formado por el producto de $I_1 = [a, b]$ e $I_2 = [c, d]$. Tomemos las siguientes particiones P_1 y P_2 de estos intervalos

$$a = x_0 < x_1 < x_2 < x_3 < x_4 = b$$
$$c = y_0 < y_1 < y_2 = d$$

Se ve que P_1 divide a I_1 en 4 subintervalos y P_2 divide a I_2 en 2 subintervalos. Entonces la partición $P = P_1 \times P_2$ de Q dividirá a este rectángulo en (4)(2) = 8 subrectángulos (a saber $Q_1 = [a, x_1] \times [c, y_1]$, $Q_2 = [a, x_1] \times [y_1, d]$, $Q_3 = [x_1, x_2] \times [y_1, d]$, $Q_4 = [x_1, x_2] \times [c, y_1]$, $Q_5 = [x_2, x_3] \times [c, y_1]$, $Q_6 = [x_2, x_3] \times [y_1, d]$, $Q_7 = [x_3, b] \times [y_1, d]$, $Q_8 = [x_3, b] \times [c, y_1]$) como se ve esquemáticamente en la figura siguiente

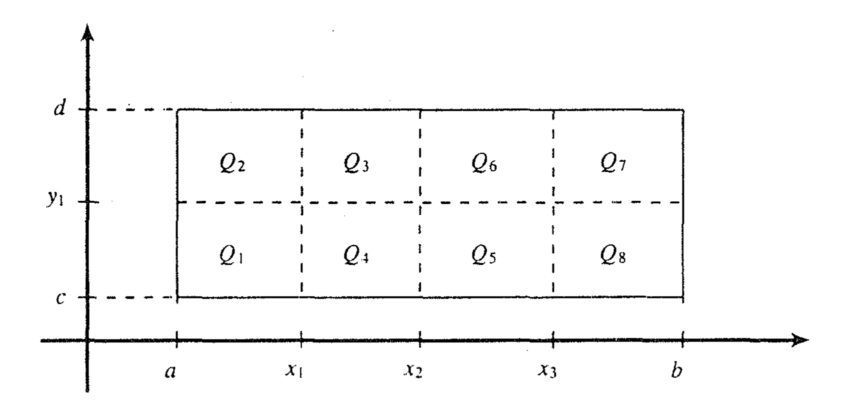


Figura 3. Una partición del rectángulo $[a, b] \times [c, d]$ en 8 subrectángulos.

Se dice que la función $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$, definida en un rectángulo Q de $\mathbb{R}^2$, es una *función escalonada*, si hay una partición P de Q, digamos que en nm subrectángulos Q_{ij}, $1 \leq i \leq n$, $1 \leq j \leq m$, de modo que la función f restringida al interior de cada uno de estos subrectángulos sea una función constante. Nuevamente advertimos que los valores de la función f en las fronteras u orillas de los subrectángulos Q_{ij} (los segmentos verticales $x = x_i$, $c \leq y \leq d$, y los segmentos horizontales $y = y_j$, $a \leq x \leq b$) son irrelevantes. Entonces, llamando f_{ij} a la restricción de f al interior del subrectángulo Q_{ij} se tiene que $f_{ij} = f|_{Q_{ij}} = c_{ij}$. La gráfica de f_{ij} es entonces, la porción del plano $z = c_{ij}$ (el cual es paralelo al plano xy) que se encuentra sobre el rectángulo Q_{ij}: así, la gráfica de la función f se puede ver como un conjunto de copias de los (interiores de los) rectángulos Q_{ij} separada cada una de ellas en forma paralela al plano xy una distancia c_{ij} (respectivamente).

Estamos ahora en condiciones de dar la definición de integral doble para una función escalonada.

Definición. Sea $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ una función escalonada definida en el rectángulo Q de $\mathbb{R}^2$. Digamos que Q está dividido en nm subrectángulos $Q_{ij} = [x_i, x_{i-1}] \times [y_j, y_{j-1}]$, y que $f_{ij}(x, y) = c_{ij}$, $1 \leq i \leq n$, $1 \leq j \leq m$. Se define la integral doble de la función $f(x, y)$ sobre el rectángulo Q, la cual denotaremos por $\iint_Q f$ o bien por $\iint_Q f(x, y)\,dx\,dy$, como

$$\iint_Q f(x, y)\,dx\,dy = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}(x_i - x_{i-1})(y_j - y_{j-1})$$

Nótese que una función *constante* $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ definida en el rectángulo Q, es una función escalonada. En tal caso, si $f(x, y) = k$ para $(x, y) \in Q = [a, b] \times [c, d]$ se tiene

$$\iint_Q f(x, y)\,dx\,dy = k(b - a)(d - c) = k \text{ (área de } Q)$$

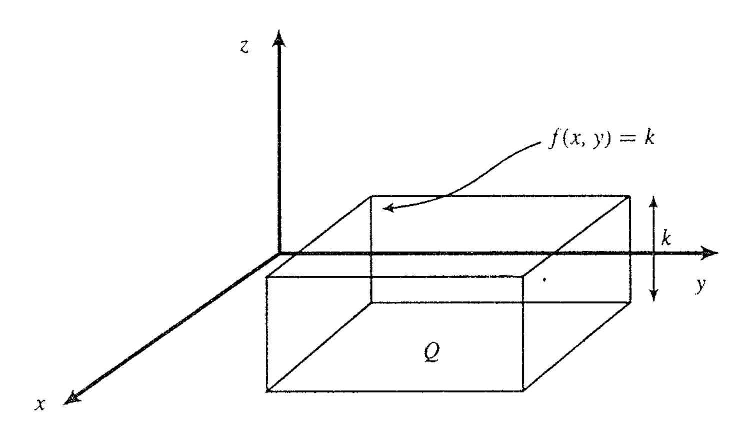


Figura 4. La integral doble de la función constante $f(x, y) = k$ sobre Q

Si $k > 0$, el valor de esta integral se puede interpretar geométricamente como el volumen de un paralelepípedo rectangular con base Q y altura k.

En realidad, podemos escribir la definición de la integral doble de la función escalonada $f(x, y)$ sobre el rectángulo Q como

$$\iint_Q f(x, y)\,dx\,dy = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}(x_i - x_{i-1})(y_j - y_{j-1}) = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}\,(\text{área de } Q_{ij})$$

donde Q_{ij} son los nm subrectángulos en que quedó dividido el rectángulo Q, sobre los cuales $f(x, y)$ es constante e igual a c_{ij}. Si todos estos valores son positivos, podemos interpretar el valor de la integral doble de $f(x, y)$ sobre Q como la suma de volúmenes de los paralelepípedos rectangulares de bases Q_{ij} y alturas c_{ij} (respectivamente) que se distinguen "debajo de la gráfica de $f(x, y)$".

Sean P_1 y P_2 dos particiones de un rectángulo Q en $\mathbb{R}^2$. Diremos que la partición P_1 es *más fina* que la partición P_2, o bien que P_1 es un refinamiento de P_2, si $P_2 \subset P_1$. (Intuitivamente podemos pensar que P_1 divide en "más pedacitos" a los ya "pedacitos" en que la partición P_2 divide a Q; por ejemplo, si en la figura 3 se añade un nuevo punto a P_1, digamos $x^* = \frac{x_2+x_3}{2}$, entonces la nueva partición $P = P_1^* \times P_2$, en que P_1^* es la partición $\{a < x_1 < x_2 < x^* < x_3 < b\}$, será un refinamiento de P: en este caso los subrectángulos Q_5 y Q_6 de la partición P quedan partidos a la mitad en la partición P^*). Por ejemplo, si P_1 y P_2 son dos particiones cualesquiera de Q, entonces su unión $P_1 \cup P_2$ es un refinamiento tanto de P_1 como de P_2.

Sea P una partición del rectángulo Q y sea P' un refinamiento de P. El lector debe convencerse de que el valor de la integral $\iint_Q f(x, y)\,dx\,dy$ para una función escalonada sobre Q, $f(x, y)$, es independiente de que en Q se tome la partición P o su refinamiento P'.

Otro hecho importante sobre el que llamamos la atención es el siguiente: si las funciones $f\colon Q \subset \mathbb{R}^2 \to \mathbb{R}$, $g\colon Q \subset \mathbb{R}^2 \to \mathbb{R}$ son dos funciones escalonadas definidas en el rectángulo Q, y α y β son dos reales cualesquiera, entonces la función $\alpha f + \beta g\colon Q \subset \mathbb{R}^2 \to \mathbb{R}$, $(\alpha f + \beta g)(x, y) = \alpha f(x, y) + \beta g(x, y)$, será también una función escalonada sobre Q. En efecto: sea P_1 la partición de Q para la cual la función $f(x, y)$ es constante en cada uno de los subrectángulos que P_1 produce en Q y sea P_2 la partición de Q para la cual $g(x, y)$ es constante en cada uno de los subrectángulos correspondientes a esta partición. La unión $P_1 \cup P_2$ es una nueva partición de Q (es, de hecho, un

refinamiento de P_1 y P_2) para la cual la función $\alpha f(x, y) + \beta g(x, y)$ es constante en cada uno de los subrectángulos de Q que tal partición produce. Es entonces una función escalonada en Q, como se quería ver.

Según la observación hecha en el párrafo anterior, tiene sentido hablar de la integral doble de la función $\alpha f(x, y) + \beta g(x, y)$ sobre el rectángulo Q. Tal como ocurre con las integrales de funciones de una sola variable, la integral doble presenta un comportamiento lineal sobre sus integrandos (sobre las funciones que se integran). Esta propiedad, junto con otras, se establece de modo formal en el teorema siguiente.

Teorema 6.1.1 Sean $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$, $g: Q \subset \mathbb{R}^2 \to \mathbb{R}$ dos funciones escalonadas definidas en el rectángulo Q.

a. Si α y β son dos números reales cualesquiera, entonces

$$\iint_Q (\alpha f(x, y) + \beta g(x, y))\, dx\, dy = \alpha \iint_Q f(x, y) dx\, dy + \beta \iint_Q g(x, y)\, dx\, dy$$

b. Si $f(x, y) \geq g(x, y)$ para toda $(x, y) \in Q$, entonces

$$\iint_Q f(x, y) dx\, dy \geq \iint_Q g(x, y) dx\, dy$$

c. Si el rectángulo Q está dividido en dos subrectángulos Q_1 y Q_2 (es decir, si se tiene $Q = Q_1 \cup Q_2$), entonces

$$\iint_Q f(x, y)\, dx\, dy = \iint_{Q_1} f(x, y)\, dx\, dy + \iint_{Q_2} f(x, y)\, dx\, dy$$ ■

La demostración de este teorema se hace directamente con la definición dada de integral doble. Resulta, sin embargo, un poco engorroso el manejo de las sumatorias que aparecen y entonces, en lugar de dar la demostración completa de este teorema, solamente presentamos un argumento breve que hace *plausible* el resultado establecido en el inciso **a**:

Sea $Q = [a, b] \times [c, d]$ (como en la figura 1). Tome en Q la partición

$$P_1 = \{a < x^* < b\} \times \{c < d\} \qquad \text{en donde } x^* = \frac{a+b}{2}$$

Sean $Q_{11} = [a, x^*] \times [c, d]$, $Q_{12} = [x^*, b] \times [c, d]$ los subrectángulos de Q en que éste queda dividido por la partición P_1. Defina en Q la función f de la manera siguiente

$$f(x, y) = \begin{cases} c_1 & \text{si } (x, y) \in \text{Int } Q_{11} \\ c_2 & \text{si } (x, y) \in \text{Int } Q_{12} \end{cases}$$

(donde Int Q significa el interior del rectángulo Q). En esquema se tiene (figura 5)

Es claro que

$$\iint_Q f(x, y)\, dx\, dy = c_1(\text{ área de } Q_{11}) + c_2(\text{ área de } Q_{12})$$

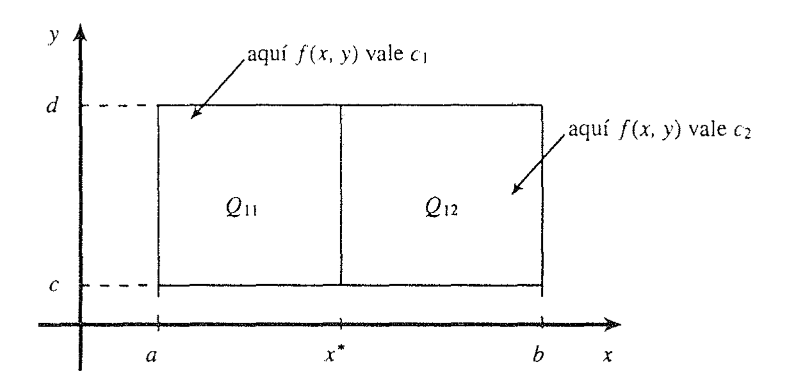


Figura 5. La función escalonada $f(x, y)$ en el rectángulo Q

Tome ahora en Q la partición

$$P_2 = \{a < b\} \times \{c < y^* < d\} \qquad y^* = \frac{c+d}{2}$$

Sean $Q_{21} = [a, b] \times [c, y^*]$ y $Q_{22} = [a, b] \times [y^*, d]$ los subrectángulos en que se divide Q por la partición P_2. Defina en Q la función g como

$$g(x, y) = \begin{cases} d_1 & \text{si } (x, y) \in \text{Int } Q_{21} \\ d_2 & \text{si } (x, y) \in \text{Int } Q_{22} \end{cases}$$

En esquema se tiene

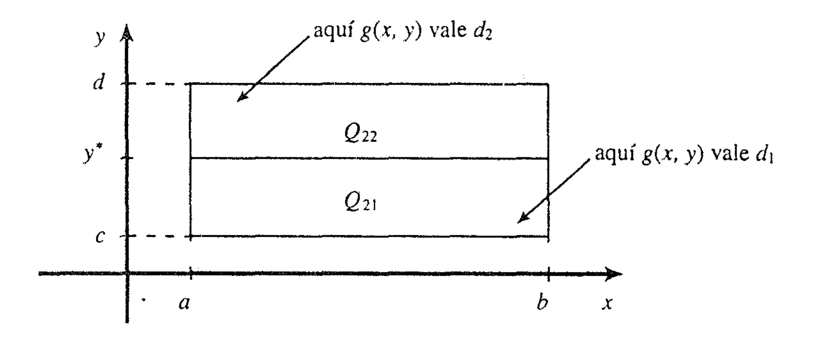


Figura 6. La función escalonada $g(x, y)$ en el rectángulo Q.

Es claro que

$$\iint_Q g(x, y)\, dx\, dy = d_1(\text{ área de } Q_{21}) + d_2(\text{ área de } Q_{22})$$

La función escalonada $\xi(x, y) = \alpha f(x, y) + \beta g(x, y)$ estará definida en el rectángulo Q con la partición $P = P_1 \bigcup P_2$, es decir

$$P = \{a < x^* < b\} \times \{c < y^* < d\}$$

Esta partición divide a Q en 4 subrectángulos, a saber

$$Q_1 = [a, x^*] \times [c, y^*], \quad Q_2 = [a, x^*] \times [y^*, d]$$
$$Q_3 = [x^*, b] \times [c, y^*], \quad Q_4 = [x^*, b] \times [y^*, d]$$

(obsérvese que $Q_{11} = Q_1 \cup Q_2$, $Q_{12} = Q_3 \cup Q_4$, $Q_{21} = Q_1 \cup Q_3$, $Q_{22} = Q_2 \cup Q_4$) y la función $\xi(x, y) = \alpha f(x, y) + \beta g(x, y)$ queda entonces como

$$\xi(x, y) = \begin{cases} \alpha c_1 + \beta d_1 & \text{si } (x, y) \in \text{Int } Q_1 \\ \alpha c_1 + \beta d_2 & \text{si } (x, y) \in \text{Int } Q_2 \\ \alpha c_2 + \beta d_1 & \text{si } (x, y) \in \text{Int } Q_3 \\ \alpha c_2 + \beta d_2 & \text{si } (x, y) \in \text{Int } Q_4 \end{cases}$$

Esquema

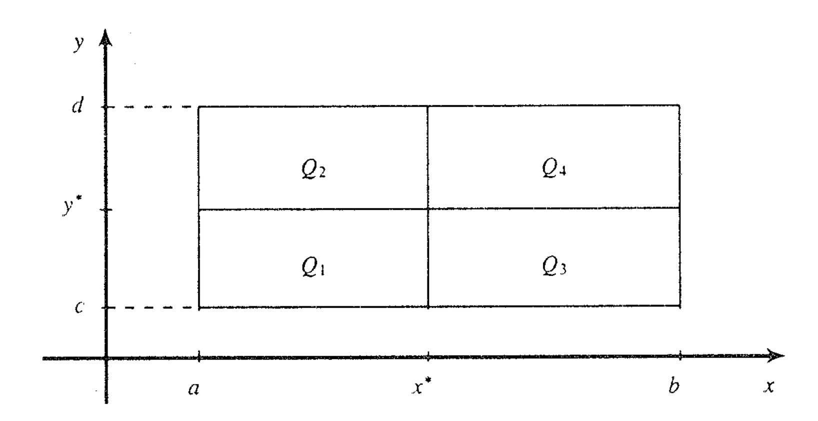


Figura 7. La función escalonada $\xi(x, y)$ en el rectángulo Q.

Se tiene entonces, atendiendo a la definición de integral doble de $\xi(x, y)$ sobre el rectángulo Q, que

$$\begin{aligned}\iint_Q (\alpha f(x, y) + \beta g(x, y))\, dx\, dy &= \iint_Q \xi(x, y)\, dx\, dy \\ &= (\alpha c_1 + \beta d_1)(\text{área de } Q_1) + (\alpha c_1 + \beta d_2)(\text{área de } Q_2) \\ &\quad + (\alpha c_2 + \beta d_1)(\text{área de } Q_3) + (\alpha c_2 + \beta d_2)(\text{área de } Q_4) \\ &= \alpha c_1(\text{área de } Q_1 + \text{ área de } Q_2) \\ &\quad + \alpha c_2(\text{área de } Q_3 + \text{área de } Q_4) \\ &\quad + \beta d_1(\text{área de } Q_1 + \text{área de } Q_3) \\ &\quad + \beta d_2(\text{ área de } Q_2 + \text{área de } Q_4) \\ &= \alpha c_1(\text{área de } Q_{11}) + \alpha c_2(\text{área de } Q_{12})\end{aligned}$$

$$+ \beta d_1(\text{ área de } Q_{21}) + \beta d_2(\text{ área de } Q_{22})$$
$$= \alpha \iint_Q f(x, y)dx\,dy + \beta \iint_Q g(x, y)dx\,dy$$

tal como asegura el inciso **a** del teorema 6.1.1. que tenía que ocurrir.

Terminamos esta sección resaltando el siguiente hecho: se observó anteriormente que si la función $f(x, y)$ definida en el rectángulo Q es una función constante, digamos $f(x, y) = k$, entonces

$$\iint_Q f(x, y)dx\,dy = k(b - a)(d - c)$$

Notando que

$$\int_a^b f(x, y)\,dx = k(b - a) \qquad \text{y} \qquad \int_c^d dy = d - c$$

podemos escribir que

$$\iint_Q f(x, y)\,dx\,dy = k(b - a)(d - c) = k(b - a)\int_c^d dy = \int_c^d k(b - a)\,dy$$
$$= \int_c^d \left(\int_a^b f(x, y)\,dx\right)dy$$

Un argumento completamente análogo muestra que

$$\iint_Q f(x, y)\,dx\,dy = \int_a^b \left(\int_c^d f(x, y)dy\right)dx$$

y así

$$\iint_Q f(x, y)\,dx\,dy = \int_c^d \left(\int_a^b f(x, y)dx\right)dy = \int_a^b \left(\int_c^d f(x, y)dy\right)dx$$

Esta fórmula, deducida para funciones constantes en el rectángulo Q, vale más bien para funciones escalonadas cualesquiera en Q. Dejamos al lector que presente los detalles del argumento que muestre este hecho. Más aún, uno de los resultados más importantes que se establecerán en este capítulo, es que la fórmula anterior es válida *para cualquier función* (de las que llamaremos integrables); este resultado establece que la integral doble se puede calcular por medio de dos integrales simples, sin importar "el orden de integración".

Ejercicios (Capítulo 6, Sección 1)

1. ¿Cómo son las funciones escalonadas de $\mathbb{R}$ en $\mathbb{R}$? Constate que la función signo y la función mayor entero de x son ejemplos de este tipo de funciones.

Considere el rectángulo $Q = [-1, 1] \times [-1, 1]$. Considere la partición de Q en 4 subrectángulos, dada por los ejes coordenados. Sea Q_i el rectángulo correspondiente al i-ésimo cuadrante (por ejemplo, $Q_2 = [-1, 0] \times [0, 1]$). En cada uno de los ejercicios 2–7, calcule el valor de la integral doble de la función dada sobre Q.

2. $f(x, y) = 5\ \forall(x, y) \in \text{Int } Q$

3. $f(x, y) = 0\ \forall(x, y) \in \text{Int } Q$

4. $f(x, y) = \begin{cases} 5 & \text{si } (x, y) \in \text{Int } Q_1 \cup \text{Int } Q_2 \\ 3 & \text{si } (x, y) \in \text{Int } Q_3 \cup \text{Int } Q_4 \end{cases}$

5. $f(x, y) = \begin{cases} 5 & \text{si } (x, y) \in \text{Int } Q_1 \cup \text{Int } Q_3 \\ -5 & \text{si } (x, y) \in \text{Int } Q_2 \cup \text{Int } Q_4 \end{cases}$

6. $f(x, y) = \begin{cases} 6 & \text{si } (x, y) \in \text{Int } Q_1 \cup \text{Int } Q_2 \cup \text{Int } Q_3 \\ 3 & \text{si } (x, y) \in \text{Int } Q_4 \end{cases}$

7. $f(x, y) = \begin{cases} 2 & \text{si } (x, y) \in \text{Int } Q_1 \\ 3 & \text{si } (x, y) \in \text{Int } Q_2 \\ 1 & \text{si } (x, y) \in \text{Int } Q_3 \\ 2 & \text{si } (x, y) \in \text{Int } Q_4 \end{cases}$

8. Sea $f(x, y)$ la función escalonada del ejercicio 2 y $g(x, y)$ la función escalonada del ejercicio 3. Describa la función $f + g$ en Q. Calcule la integral doble de esta función en Q.

9. Sea $f(x, y)$ la función escalonada del ejercicio 4 y $g(x, y)$ la función escalonada del ejercicio 5. Describa la función $f + g$ en Q. Calcule la integral doble de esta función en Q.

10. Sea $f(x, y)$ la función escalonada del ejercicio 6 y $g(x, y)$ la función escalonada del ejercicio 7. Describa la función $f + g$ en Q. Calcule la integral doble de esta función en Q.

Considere el rectángulo $Q = [-1, 1] \times [-1, 1]$ y la partición de él en los 3 subrectángulos $Q_1^* = [-1, -1/2] \times [-1, 1]$, $Q_2^* = [-1/2, 1/2] \times [-1, 1]$, $Q_3^* = [1/2, 1] \times [-1, 1]$. Sea $f(x, y)$ la función dada en Q con la partición descrita en los ejercicios 2–7. En los ejercicios 11–14 se dan las funciones $f(x, y)$ y $g(x, y)$ en Q. Describa en cada uno de ellos la función escalonada $\alpha f(x, y) + \beta g(x, y)$, en que α y β son los números dados, y verifique la propiedad **a** del teorema 6.1.1.

11. $f(x, y)$ es la función del ejercicio 4.
$g(x, y) = \begin{cases} 3 & \text{si } (x, y) \in \text{Int } Q_1^* \cup \text{Int } Q_2^* \\ 5 & \text{si } (x, y) \in \text{Int } Q_3^* \end{cases}$
$\alpha = -2, \beta = 2$

12. $f(x, y)$ es la función del ejercicio 5.
$g(x, y) = \begin{cases} 2 & \text{si } (x, y) \in \text{Int } Q_1^* \\ 1 & \text{si } (x, y) \in \text{Int } Q_2^* \cup \text{Int } Q_3^* \end{cases}$
$\alpha = -3, \beta = 7$

13. $f(x, y)$ es la función del ejercicio 6.
$g(x, y) = \begin{cases} 1 & \text{si } (x, y) \in \text{Int } Q_1^* \\ 0 & \text{si } (x, y) \in \text{Int } Q_2^* \\ 8 & \text{si } (x, y) \in \text{Int } Q_3^* \end{cases}$
$\alpha = 1, \beta = -1$

14. $f(x, y)$ es la función del ejercicio 7.

$$g(x, y) = \begin{cases} 3 & \text{si } (x, y) \in \text{Int } Q_1^* \\ 5 & \text{si } (x, y) \in \text{Int } Q_2^* \\ 7 & \text{si } (x, y) \in \text{Int } Q_3^* \end{cases}$$

$\alpha = -4, \beta = 8$

6.2 Integrales dobles (II): funciones integrables sobre rectángulos

En esta sección presentamos una definición rigurosa de integral doble de una función $f(x, y)$ definida en un rectángulo Q de $\mathbb{R}^2$. La clase de funciones que vamos a considerar para esta definición es la de las funciones *acotadas*: recuerde que una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ se dice ser acotada (en U) si se da una constante $M > 0$ de modo que $|f(x, y)| \leq M, \forall(x, y) \in U$. La idea para establecer el concepto de integral doble de una función acotada $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ definida en el rectángulo Q es atrapar a esta función por encima y por debajo con funciones escalonadas, acerca de las cuales ya se definió en la sección anterior el concepto de integral doble sobre Q, y considerar las integrales de todas estas funciones. Si "nos queda algo en medio" de todas las integrales dobles (sobre Q) de las funciones escalonadas que están por debajo y por encima de $f(x, y)$, a este algo es a lo que llámaremos integral doble de $f(x, y)$ en Q. Para proceder de esta manera, es necesario primero convencerse de que dichas funciones escalonadas (que atrapan a f), en realidad *existen*. Esto queda garantizado por el hecho de que f es acotada. En efecto, puesto que $|f(x, y)| \leq M, \forall(x, y) \in Q$, se tiene que $-M \leq f(x, y) \leq M, \forall(x, y) \in Q$. Las funciones $\phi, \psi: Q \to \mathbb{R}$ definidas como $\phi(x, y) = -M, \psi(x, y) = M$ son funciones escalonadas (de hecho constantes) sobre Q, y "atrapan" a f por debajo y por encima, pues $\phi(x, y) \leq f(x, y) \leq \psi(x, y), \forall(x, y) \in Q$.

Establezcamos entonces la definición correspondiente.

Definición. Sea $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ una función acotada definida en el rectángulo Q de $\mathbb{R}^2$. Si hay un único número real I, de modo que

$$\iint_Q \phi(x, y)\, dx\, dy \leq I \leq \iint_Q \psi(x, y)\, dx\, dy$$

para todas las funciones escalonadas $\phi, \psi: Q \subset \mathbb{R}^2 \to \mathbb{R}$ según lo cual

$$\phi(x, y) \leq f(x, y) \leq \psi(x, y) \qquad \forall(x, y) \in Q$$

se dice que la función f es *integrable* sobre Q, entonces al número I se le llama integral doble de f sobre Q y se escribe

$$I = \iint_Q f(x, y) dx\, dy$$

Como ya habíamos mencionado en la introducción del capítulo, la definición anterior presenta de una manera sumamente decente desde el punto de vista matemático lo que es una integral doble de una función acotada en un rectángulo. Sin embargo, resulta muy poco práctica para *calcular* integrales de este tipo de funciones (resulta poco práctica hasta para poner un ejemplo concreto que ilustre la

existencia de la integral doble de una de dichas funciones). El teorema siguiente nos empieza a dar luz en cuanto al cálculo práctico de estas integrales. Es un resultado de suma importancia que nos dice que, en efecto, una integral doble se puede ver, como su nombre lo sugiere, como una operación que se realiza sobre la función correspondiente calculando *dos integrales* (de las que conocemos de nuestro primer curso de cálculo).

Teorema 6.2.1 Sea $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ una función acotada e integrable en el rectángulo $Q = [a, b] \times [c, d]$. Supongamos que para cada $y \in [c, d]$, exista la función $g: [c, d] \to \mathbb{R}$, $g(y) = \int_a^b f(x, y)\,dx$ (es decir, la función $f(x, y)$ vista como función de la variable x es integrable). Si la función g es integrable en $[c, d]$, su integral es igual a la integral doble de $f(x, y)$ sobre Q, con lo que se tiene

$$\iint_Q f(x, y)\,dx\,dy = \int_c^d g(y)\,dy = \int_c^d \left(\int_a^b f(x, y)\,dx\right) dy$$

De igual modo, si existe para cada $x \in [a, b]$, la función $h: [a, b] \to \mathbb{R}$ definida como $h(x) = \int_c^d f(x, y)dy$ (es decir, la función $f(x, y)$ vista como función de la variable y es integrable); si esta función $h(x)$ es integrable en $[a, b]$, su integral es la integral doble de $f(x, y)$ en Q

$$\iint_Q f(x, y)\,dx\,dy = \int_a^b h(x)\,dx = \int_a^b \left(\int_c^d f(x, y)\,dy\right) dx$$

Demostración. Sean $\phi(x, y)$ y $\psi(x, y)$ dos funciones escalonadas en Q tales que $\phi(x, y) \leq f(x, y) \leq \psi(x, y)$, $\forall (x, y) \in Q$. Ciertamente las funciones escalonadas son funciones que, vistas con una de sus variables fijas, dependiendo solamente de la otra variable, son integrables (pues vistas así resultan ser las funciones escalonadas que se estudian en los cursos de cálculo de una sola variable, las cuales sabemos que son integrables). Viendo entonces la desigualdad anterior como una desigualdad entre tres funciones integrables que dependen de la variable x (con la y fija), podemos integrar cada una de ellas en el intervalo $[c, d]$ y conservar el sentido de las desigualdades quedándonos

$$\int_a^b \phi(x, y)\,dx \leq \int_a^b f(x, y)\,dx \leq \int_a^b \psi(x, y)\,dx$$

o sea

$$\int_a^b \phi(x, y)\,dx \leq g(y) \leq \int_a^b \psi(x, y)\,dx$$

Siendo $g(y)$ integrable, al igual que las funciones de la variable y que quedaron en los extremos de esta desigualdad (¿por qué?; ver la última fórmula establecida en la sección anterior), podemos integrar respecto de y en el intervalo $[c, d]$ para que nos quede

$$\int_c^d \left(\int_a^b \phi(x, y)dx\right) dy \leq \int_c^d g(y)\,dy \leq \int_c^d \left(\int_a^b \psi(x, y)\,dx\right) dy$$

o sea

$$\iint_Q \phi(x, y)\,dx\,dy \leq \int_c^d g(y)\,dy \leq \iint_Q \psi(x, y)\,dx\,dy$$

Puesto que las funciones escalonadas que tomamos son arbitrarias, y la función $f(x, y)$ es integrable en Q, el número $I = \int_c^d g(y)\,dy$ debe ser la integral doble de $f(x, y)$ sobre Q. Esto es lo que se quería probar. El argumento que prueba la segunda parte del teorema es completamente similar y dejamos al lector que escriba los detalles del mismo. Q.E.D.

Las hipótesis hechas sobre la función $f(x, y)$ en el teorema anterior que permiten concluir su integrabilidad en el rectángulo Q podrían parecer muy restrictivas. En realidad no lo son tanto. Para finalizar esta primera parte teórica del capítulo, enunciaremos un teorema referente a una gran clase de funciones que cumplen las condiciones establecidas en el teorema estudiado. Tales funciones son las *continuas*. En realidad, se puede probar que las funciones $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$, aun sin ser continuas en el rectángulo Q, pueden ser integrables y valer las fórmulas establecidas en el teorema, con tal de que "no sean muchos los puntos en que tales funciones sean discontinuas". Este es un resultado que tendremos que usar para poder estudiar integrales dobles de funciones sobre regiones en $\mathbb{R}^2$ más generales. En el apéndice de esta sección estableceremos las definiciones correspondientes y la demostración del teorema que enuncia con precisión el resultado. Por ahora vamos a enunciar el teorema mencionado que dice que "funciones continuas son integrables". Su demostración es muy técnica y la omitiremos (puede consultarse, por ejemplo, en [ApII], pag. 443).

Teorema 6.2.2 Si la función $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ definida en el rectángulo $Q = [a, b] \times [c, d]$ es continua, entonces es integrable, y la integral doble de ella sobre Q se puede calcular como

$$\iint_Q f(x, y)\,dx\,dy = \int_c^d \left(\int_a^b f(x, y)\,dx \right) dy = \int_a^b \left(\int_c^d f(x, y)dy \right) dx \qquad ■$$

NOTA: La diferencia entre las dos fórmulas que conecta la segunda igualdad en la expresión anterior es *el orden de integración*: en la primera se integra primero respecto de x y luego respecto de y, y en la segunda es al contrario. El hecho de que se use una u otra fórmula al calcular la integral doble depende de la estructura de la función $f(x, y)$, que puede facilitar el cálculo de alguna de las dos integrales procediendo con un orden determinado. En otros casos, sin embargo, el orden de integración resulta ser irrelevante.

Veamos algunos ejemplos concretos de cálculo de integrales dobles de funciones continuas definidas en rectángulos.

Ejemplo 1. Calculemos la integral doble $\iint_Q (x^2+y^2)\,dx\,dy$ sobre el rectángulo $Q = [0, 2]\times[1, 2]$. Es conveniente tener la imagen geométrica de nuestra región de integración (figura 1).

Así pues, integraremos de 0 a 2 respecto de x y de 1 a 2 respecto de y. La integral doble queda entonces como

$$\iint_Q (x^2 + y^2)dx\,dy = \int_1^2 \left(\int_0^2 (x^2 + y^2)\,dx \right) dy =$$

$$= \int_1^2 \left(\int_0^2 x^2 dx + \int_0^2 y^2 dx \right) dy = \int_1^2 \left(\frac{1}{3}x^3 + y^2 x \right)\Big|_0^2 dy$$

$$= \int_1^2 \left(\frac{8}{3} + 2y^2 \right) dy = \left(\frac{8}{3}y + \frac{2}{3}y^3 \right)\Big|_1^2 = \frac{8}{3}(2 - 1) + \frac{2}{3}(8 - 1) =$$

$$= \frac{22}{3} \qquad ■$$

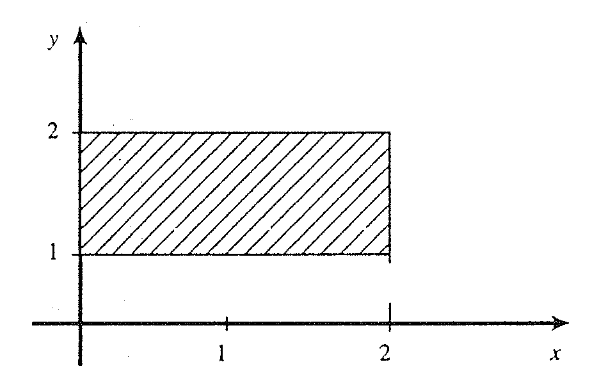


Figura 1. Región de integración para el ejemplo 1

Ejemplo 2. Calculemos la integral de la función $f(x, y) = \frac{xy}{1+x^2+y^2}$ en el rectángulo $Q = [-1, 1] \times [-1, 1]$. Tanto en x como en y, integraremos de -1 a 1. Se tiene entonces

$$\iint\limits_Q \frac{xy}{1+x^2+y^2} dx\, dy = \int_0^1 \left(\int_0^1 \frac{xy}{1+x^2+y^2} dy \right) dx$$

$$= \int_0^1 \left(\frac{x}{2} \ln(1 + x^2 + y^2) \right)_0^1 dx$$

(en esta integral lo que evaluamos de 0 a 1 es la y, pues fue la variable respecto a la que se integró)

$$= \int_0^1 \left(\frac{x}{2} \ln(2 + x^2) - \frac{x}{2} \ln(1 + x^2) \right) dx$$

$$= \frac{1}{4} \left((2 + x^2)\big(\ln(2 + x^2) - 1\big) - (1 + x^2)\big(\ln(1 + x^2) - 1\big) \right)_0^1$$

$$= \frac{1}{4} \big(3(\ln 3 - 1) - 2(\ln 2 - 1) - 2(\ln - 1) - 1\big)$$

$$= \frac{1}{4} \ln \frac{27}{16}$$ ■

Es interesante dar una interpretación geométrica al teorema 6.2.1. Trataremos de presentar argumentos que, sin ser rigurosos, hagan plausible la conclusión a la que se desea llegar. En la sección anterior se pudo concluir que para una función constante $f(x, y) = c$ definida en el rectángulo Q, su integral doble es el volumen del paralelepípedo rectangular con Q y altura c. No resultará extraño que la conclusión sea que la integral doble de la función $f(x, y)$ sobre Q es también un volumen de "un paralelelpípedo rectangular cuya tapa es la gráfica de la función $f(x, y)$ sobre Q". Consideremos pues una función acotada $f: Q \subset \mathbb{R}^2 \to \mathbb{R}$ definida en el rectángulo $Q = [a, b] \times [c, d]$ de modo que $f(x, y) \geq 0\ \forall (x, y) \in Q$ (es decir, la función f es no negativa). Entonces la superficie que representa la gráfica de f nunca está por debajo del plano xy dentro del rectángulo Q. Consideremos el cuerpo Ω que queda limitado entre la gráfica de $f(x, y)$, el plano xy, y el "cilindro" limitado por el rectángulo Q. Para cada $y \in [c, d]$, la función $g: [c, d] \to \mathbb{R}$,

$g(y) = \int_a^b f(x, y)\,dx$ nos da el área de la sección transversal que se obtiene de la intersección del cuerpo Ω con el plano perpendicular al plano xy correspondiente al valor considerado de y (del tipo $y =$ constante). De esta manera, el cuerpo Ω queda (infinitamente) dividido en rebanadas paralelas al plano xz, con el área de cada una de ellas dada por el valor de $g(y)$. Siendo esta función integrable y pensando de manera intuitiva que al integrar esta función desde $y = c$ hasta $y = d$, lo que haremos será pegar las infinitas rebanadas (infinitamente delgadas); no debe resultar difícil aceptar que lo que finalmente se obtiene es el *volumen de* Ω. Así pues se tiene que

$$\iint_Q f(x, y)\,dx\,dy = \int_c^d g(y)\,dy = \text{Volumen de } \Omega$$

El argumento anterior es el mismo si consideramos las rebanadas de Ω que se obtienen al cortar con planos paralelos al plano yz (del tipo $x =$ constante), cuyas áreas están dadas por $h(x) = \int_c^d f(x, y)\,dy$. Al pegar todas estas rebanadas, integrando $h(x)$ desde $x = a$ hasta $x = b$ obtendremos el volumen de Ω. Es decir

$$\iint_Q f(x, y)\,dx\,dy = \int_a^b h(x)\,dx = \text{Volumen de } \Omega$$

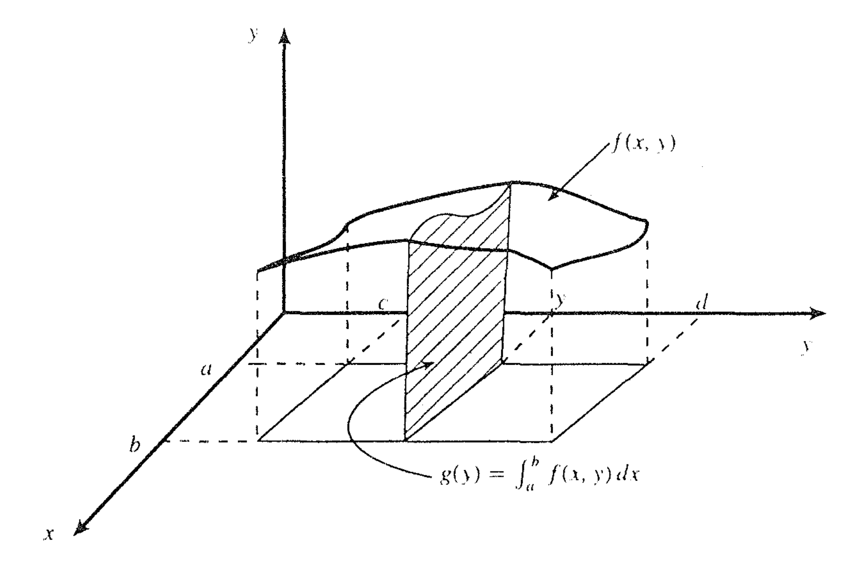


Figura 2. Interpretación geométrica del teorema 6.2.1.

El apéndice siguiente contiene las definiciones y resultados rigurosos que nos permitirán establecer un teorema importante sobre integrabilidad de funciones no continuas. Este resultado se usará en la siguiente sección, en la que se considera el problema de integración de funciones sobre regiones más generales (no rectangulares). Sin embargo, advertimos que el material aquí presentado es de carácter opcional, y el lector puede continuar a la siguiente sección, recurriendo (sólo en su momento) a leer (y entender) los resultados de referencia.

Apéndice Integrabilidad de funciones discontinuas en conjuntos de medida cero

Sea C un conjunto acotado del plano $\mathbb{R}^2$. Diremos que C tiene *medida cero* si dado $\epsilon > 0$ se da un conjunto finito de rectángulos Q_i, $i = 1, 2, \ldots, n$ de modo que la unión de éstos contenga a C, y la suma de sus áreas sea menor o igual que ϵ. En otras palabras: el conjunto C es de medida cero si lo podemos "cubrir" con una cantidad finita de rectángulos cuya suma de áreas es menor o igual que cualquier número positivo prefijado. La idea intuitiva que hay detrás de este concepto es que los conjuntos de medida cero (en el plano) son "conjuntos flacos" en el sentido de que "su área es cero" Veamos algunos ejemplos.

Ejemplo 3. Sea C un conjunto finito de puntos en el plano, digamos $\mathbf{p}_i = (x_i, y_i)$, $i = 1, 2, \ldots, n$. Este es un conjunto de medida cero. En efecto, es claro que podemos construir un rectángulo Q_i con centro en $\mathbf{p}_i$ cuya área sea menor o igual a ϵ/n (dejamos como ejercicio para el lector que dé explícitamente los vértices de dicho rectángulo). Si Q es la unión de los Q_i, entonces Q contiene a C y la suma de las áreas de los rectángulos de Q es menor o igual a ϵ. ■

Ejemplo 4. Si $C_1, C_2, \ldots, C_k$ son conjuntos de medida cero, entonces el conjunto $C = C_1 \cup C_2 \cup \ldots \cup C_k$ también tiene medida cero. En efecto, el conjunto C_j puede ser cubierto con una reunión R_j de rectángulos cuya suma de áreas no supera a ϵ/k, $j = 1, 2, \ldots, k$. Sea Q la unión de los R_j. Entonces es claro que Q contiene a C y la suma de las áreas de los rectángulos de Q no supera a ϵ ■

Un ejemplo muy importante de conjunto de medida cero es el de la gráfica de una función $\phi(x)$ continua en el intervalo $[a, b]$. El teorema es el siguiente. (Para entender su demostración hay que saber algo acerca de continuidad uniforme. Si al lector le suenan extraños estos términos, no se preocupe; haga caso omiso de ella y trate de entender solamente lo que dice el teorema)

Teorema 6.2.3 Sea $\phi: [a, b] \to \mathbb{R}$ una función continua. Entonces la gráfica de ϕ (es decir, el conjunto $C = \{(x, y) \in \mathbb{R}^2 | x \in [a, b], y = \phi(x)\}$) tiene medida cero. ■

Demostración. Siendo ϕ una función continua en el intervalo compacto $[a, b]$, es uniformemente continua en dicho intervalo. Es decir, dado $\epsilon > 0$ existe $\delta > 0$ de modo que

$$x, y \in [a, b], \ |x - y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{b - a}$$

Sea $n \in \mathbb{N}$ tal que $\frac{b-a}{n} < \delta$. Tome una partición del intervalo $[a, b]$ en n partes iguales. Es decir, considere la partición P de $[a, b]$ dada por

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b, \quad \text{en donde } x_i = a + i\frac{b - a}{n}$$

Se tiene entonces que

$$x, y \in [x_{i-1}, x_i] \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{b - a}$$

Nótese que en cada subintervalo $[x_{i-1}, x_i]$, la gráfica de ϕ queda dentro de un rectángulo Q_i de altura $\frac{\epsilon}{b-a}$. Así, la gráfica de la función ϕ queda contenida en la unión de estos n rectángulos Q_i cuya suma de áreas es

$$\sum_{i=1}^{n}(x_i - x_{i-1})\frac{\epsilon}{b-a} = \frac{\epsilon}{b-a}\sum_{i=1}^{n}(x_i - x_{i-1}) = \frac{\epsilon}{b-a}(b-a) = \epsilon$$

Es decir, la gráfica de ϕ tiene medida cero, como se quería ver. Q.E.D.

En el teorema 6.2.2 vimos que una función $f: Q \to \mathbb{R}$ definida en el rectángulo $Q = [a, b] \times [c, d]$, *continua*, es integrable. El siguiente resultado, el más importante de este apéndice, que usaremos en la siguiente sección, debilita la hipótesis de continuidad de la función f en Q: aun cuando esta función sea discontinua, podemos concluir su integrabilidad en Q, con tal de que su cantidad de discontinuidades no sea muy grande. Esta condición de la cantidad de discontinuidades se establece en forma rigurosa diciendo que el conjunto de discontinuidades de f sea de medida cero. Enunciamos ahora con precisión el teorema correspondiente. Omitimos la demostración la (puede consultarse, por ejemplo, en [ApII], pag. 446).

Teorema 6.2.4 Sea $f: Q \to \mathbb{R}$ una función definida y acotada en el rectángulo $Q = [a, b] \times [c, d]$. Si el conjunto de discontinuidades de f en Q tiene medida cero, entonces f es integrable en Q. ■

Ejemplo 5. Sea $Q = [-1, 1] \times [-1, 1]$. Consideremos la función $f: Q \to \mathbb{R}$ dada por $f(x, y) = \frac{xy}{1+x^2+y^2}$ (ver ejemplo 2). Sea A un subconjunto finito de puntos de Q. Defina la función $g: Q \to \mathbb{R}$ como

$$g(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in Q - A \\ K & \text{si } (x, y) \in A \end{cases}$$

donde K es un número real dado. En general esta función es discontinua en Q. Sin embargo, el conjunto de discontinuidades A tiene medida cero (pues es un conjunto finito de puntos de Q). Entonces, por el teorema anterior, la función $g(x, y)$ es integrable en Q y su integral es, de hecho, *la misma* que la de la función $f(x, y)$ en Q (¿por qué?), es decir

$$\iint_Q g(x, y)\,dx\,dy = \iint_Q \frac{xy}{1+x^2+y^2}\,dx\,dy = \frac{1}{4}\ln\frac{27}{16}$$ ■

Ejemplo 6. Sea $Q = [-1, 1] \times [0, 1]$. Dividamos este rectángulo en dos regiones A y B definidas como

$$A = \{(x, y) \in Q | y \geq x^2\}, \quad B = \{(x, y) \in Q | y < x^2\}$$

Defina la función $f: Q \to \mathbb{R}$ como

$$f(x, y) = \begin{cases} x^2y^3 & \text{si } (x, y) \in A \\ 0 & \text{si } (x, y) \in B \end{cases}$$

En esquema tenemos

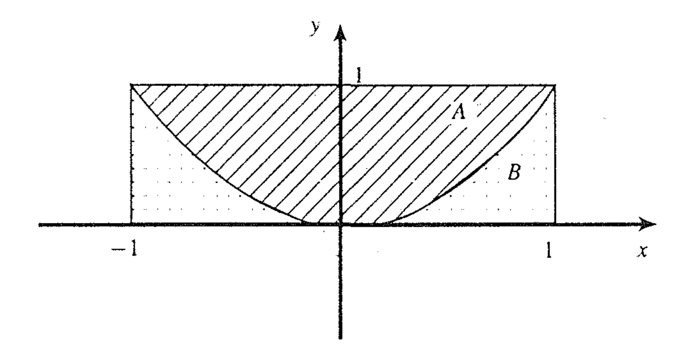


Figura 3. Región de integración para el ejemplo 5.

Obsérvese que la función $f(x, y)$ es continua en todo el rectángulo Q, *excepto en los puntos de la parábola* $y = x^2$ (menos el origen, donde sí es continua —¿por qué?—). Estos puntos de discontinuidad de $f(x, y)$ son los de la gráfica de una función continua ($y = x^2$) y por lo tanto, forman un conjunto de medida cero. Así la función $f(x, y)$ entonces integrable en Q. Calculemos su integral

$$\iint_Q f(x, y)dx\,dy = \int_{-1}^{1}\left(\int_0^1 f(x, y)dy\right)dx =$$

$$= \int_{-1}^{1}\left(\int_0^{x^2} f(x, y)dy + \int_{x^2}^{1} f(x, y)dy\right)dx =$$

$$= \int_{-1}^{1}\left(\int_0^{x^2} 0dy + \int_{x^2}^{1} x^2y^3dy\right)dx = \int_{-1}^{1}\left(\int_{x^2}^{1} x^2y^3dy\right)dx =$$

$$= \int_{-1}^{1}\left(\frac{1}{4}x^2y^4\right)\Big|_{x^2}^{1} dx = \int_{-1}^{1} \frac{1}{4}x^2(1 - x^8)dx = \frac{4}{33}$$

Si cambiamos el orden de integración para calcular la integral, los cálculos se verían como

$$\iint_Q f(x, y)\,dx\,dy = \int_0^1\left(\int_{-1}^{1} f(x, y)\,dx\right)dy$$

$$= \int_0^1\left(\int_{-1}^{-\sqrt{y}} 0\,dx + \int_{-\sqrt{y}}^{\sqrt{y}} x^2y^3\,dx + \int_{\sqrt{y}}^{1} 0dx\right)dy$$

$$= \int_0^1\left(\int_{-\sqrt{y}}^{\sqrt{y}} x^2y^3\,dx\right)dy = \int_0^1\left(\frac{1}{3}x^3y^3\right)\Big|_{-\sqrt{y}}^{\sqrt{y}} dy$$

$$= \int_0^1 \frac{1}{3}y^3(2y^{3/2})\,dy = \int_0^1 \frac{2}{3}y^{9/2}\,dy$$

$$= \left(\frac{2}{3}\frac{2}{11}y^{11/2}\right)_0^1 = \frac{4}{33}$$ ■

Ejercicios (Capítulo 6, Sección 2)

En cada uno de los ejercicios 1—12, calcule la integral doble de la función $f(x, y)$ dada sobre el rectángulo Q.

1. $f(x, y) = 3, \quad Q = [1, 3] \times [2, 3]$

2. $f(x, y) = x, \quad Q = [-2, 0] \times [2, 4]$

3. $f(x, y) = y, \quad Q = [2, 4] \times [-2, 0]$

4. $f(x, y) = 2xy, \quad Q = [-1, 0] \times [0, 1]$

5. $f(x, y) = 2x + 4y, \quad Q = [-2, 2] \times [2, 3]$

6. $f(x, y) = (x + 2y)^3, \quad Q = [-1, 5] \times [3, 7]$

7. $f(x, y) = \operatorname{sen}(x + 4y), \quad Q = [2, 5] \times [3, 6]$

8. $f(x, y) = x\cos(2x - y), \quad Q = [1, 2] \times [3, 4]$

9. $f(x, y) = e^{x+y}, \quad Q = [1, 2] \times [0, 3]$

10. $f(x, y) = e^x \operatorname{sen} y, \quad Q = [0, 1] \times [0, 1]$

11. $f(x, y) = x^2ye^{xy}, \quad Q = [0, 1] \times [0, 1]$

12. $f(x, y) = \dfrac{1}{(2x + y - 3)^3}, \quad Q = [2, 3] \times [2, 3]$

6.3 Integrales dobles de funciones sobre regiones más generales

Ampliaremos ahora el concepto de integral doble de una función $f(x, y)$ (hecho en la sección anterior), sobre rectángulos Q, a regiones más generales en $\mathbb{R}^2$. Estos serán sobconjuntos acotados del plano (que podamos incluir en un rectángulo Q) y las clasificamos en tres tipos:

REGIONES DEL TIPO (I). Sean $\phi_1, \phi_2 \colon [a, b] \to \mathbb{R}$ dos funciones reales de la variable real $x \in [a, b]$, continuas, de modo que $\phi_1(x) \leq \phi_2(x)$, $\forall x \in [a, b]$. Consideremos la región D_I del plano dada por

$$D_I = \{(x, y) | a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$

El hecho de que las funciones ϕ_1 y ϕ_2 sean funciones continuas en el intervalo cerrado $[a, b]$, garantiza que son acotadas, lo cual a su vez garantiza que la región D_I es una región acotada del plano, cuyo aspecto es como el que se muestra en la figura siguiente (figura 1).

Así pues, las regiones del tipo (I) son regiones limitadas:

(*) por la recta vertical $x = a$ por la izquierda,

(*) por la recta vertical $x = b$ por la derecha,

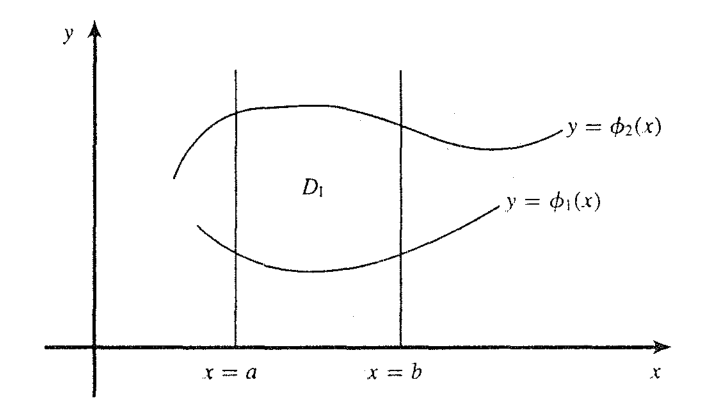


Figura 1. Una región del tipo (I) en el plano.

(*) por la gráfica de la función de x, $y = \phi_1(x)$ por debajo, y,

(*) por la gráfica de la función de x, $y = \phi_2(x)$ por encima.

REGIONES DEL TIPO (II). Sean $\psi_1, \psi_2: [c, d] \to \mathbb{R}$ dos funciones reales de la variable real $y \in [c, d]$, continuas, de modo que $\psi_1(y) \leq \psi_2(y)$, $\forall y \in [c, d]$. Consideremos la región D_{II} del plano definida por

$$D_{II} = \{(x, y) | \psi_1(y) \leq x \leq \psi_2(y), c \leq y \leq d\}$$

Nuevamente, por la continuidad de las funciones $\psi_1(y)$ y $\psi_2(y)$ en el intervalo cerrado $[c, d]$, podemos garantizar que la región D_{II} es acotada. Su aspecto gráfico es como sigue

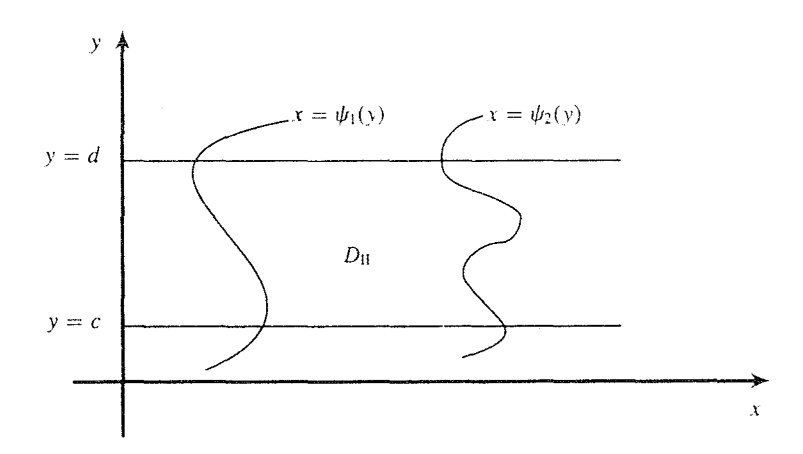


Figura 2. Una región del tipo (II) en el plano.

De modo pues que las regiones del tipo (II) están limitadas:

(*) por la recta horizontal $y = c$ por debajo,

(*) por la recta horizontal $y = d$ por encima,

(*) por la gráfica de la función de y, $x = \psi_1(y)$ por la izquierda, y,

(*) por la gráfica de la función de y, $x = \psi_2(y)$ por la derecha.

REGIONES DEL TIPO (III). Vamos también a considerar regiones que se pueden descomponer, por medio de un número finito de "cortes" verticales y/o horizontales, en regiones del tipo (I) y/o regiones del tipo (II). Casos concretos de regiones de este tipo se verán en los ejemplos que se presentan a continuación.

Ejemplo 1. Consideremos la región comprendida entre las parábolas $y = 4 - x^2$, $y = x^2 - 4$. Esta se puede ver de manera natural como una región del tipo (I). Viendo las intersecciones entre estas parábolas

$$4 - x^2 = x^2 - 4 \Rightarrow x = 2, -2$$

se tiene que la región queda limitada:

(*) por la izquierda por la recta $x = -2$,

(*) por la derecha por la recta $x = 2$,

(*) por debajo por la parábola $\phi_1(x) = x^2 - 4$,

(*) por encima por la parábola $\phi_2(x) = 4 - x^2$.

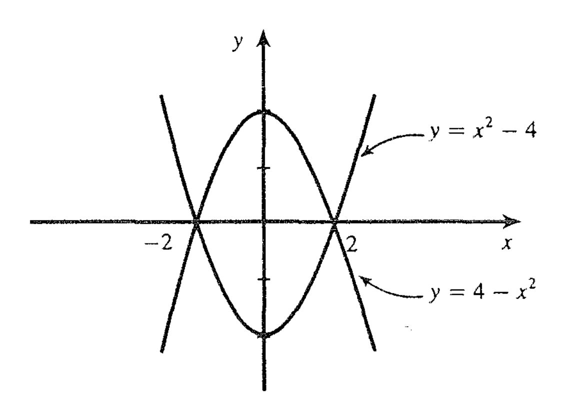


Figura 3. La región del ejemplo 1. ■

Ejemplo 2. Consideremos el trapecio cuyos vértices son $A = (1, 1)$, $B = (4, 1)$, $C = (2, 4)$ y $D = (3, 4)$ (figura 4).

La región D limitada por este trapecio se puede ver de manera natural como una región del tipo (II). En efecto, la recta ℓ_1 que pasa por A y C es $y = 3x - 2$, o bien $x = \psi_1(y) = (y+2)/3$, y la recta ℓ_2 que pasa por B y D es $y = -3x + 13$, o bien $x = \psi_2(y) = (13 - y)/3$, de modo que la región que el trapecio encierra se puede ver limitada:

(*) por la recta horizontal $y = 1$ por debajo,

(*) por la recta horizontal $y = 4$ por encima,

(*) por la recta $x = \psi_1(y) = (y + 2)/3$ por la derecha, y,

(*) por la recta $x = \psi_2(y) = (13 - y)/3$ por la izquierda. ■

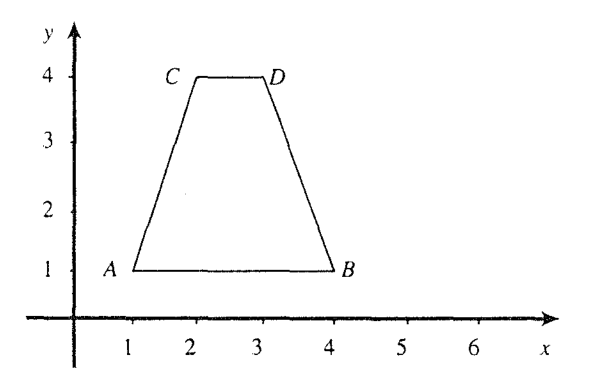


Figura 4. La región del ejemplo 2.

Ejemplo 3. Consideremos la región D limitada por el círculo unitario $x^2 + y^2 = 1$. Esta es una región que puede verse como del tipo (I) o del tipo (II). En efecto, si despejamos la y de la ecuación que define al círculo tenemos dos funciones de x, $\phi_1(x) = -\sqrt{1 - x^2}$, $\phi_2(x) = \sqrt{1 - x^2}$ que limitan a D por debajo y por encima, respectivamente, en tanto que por la izquierda y por la derecha podemos pensar en D limitada por las rectas verticales $x = -1$ y $x = 1$. Por otra parte, despejando la x de la ecuación que define al círculo obtenemos las dos funciones de y, $\psi_1(y) = -\sqrt{1 - y^2}$ y $\psi_2(y) = \sqrt{1 - y^2}$ que limitan a D por la izquierda y por la derecha, respectivamente, en tanto que D queda limitada por debajo y por arriba por las rectas $y = -1$ e $y = 1$.

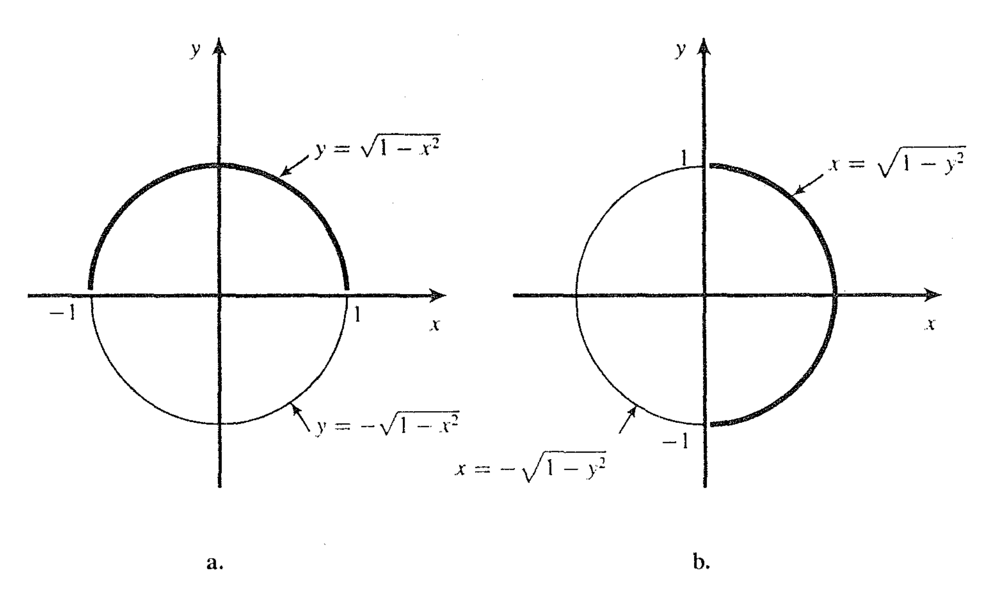


Figura 5. La región D limitada por el círculo $x^2 + y^2 = 1$, vista como: **a.** región del tipo (I), **b.** región del tipo (II).

Ejemplo 4. Consideremos el paralelogramo cuyos vértices son $A = (0, 0)$, $B = (3, 1)$, $C = (4, 4)$ y $D = (1, 3)$. Las rectas que unen cada par de estos cuatro puntos son:

a. la que pasa por A y B, $y = \phi_1(x) = x/3$, o bien, $x = \psi_1(y) = 3y$.

b. la que pasa por B y C, $y = \phi_2(x) = 3x - 8$, o bien, $x = \psi_2(y) = (y + 8)/3$

c. la que pasa por C y D, $y = \phi_3(x) = (x + 8)/3$, o bien, $x = \psi_3(y) = 3y - 8$.

d. la que pasa por D y A, $y = \phi_4(x) = 3x$, o bien, $x = \psi_4(y) = y/3$.

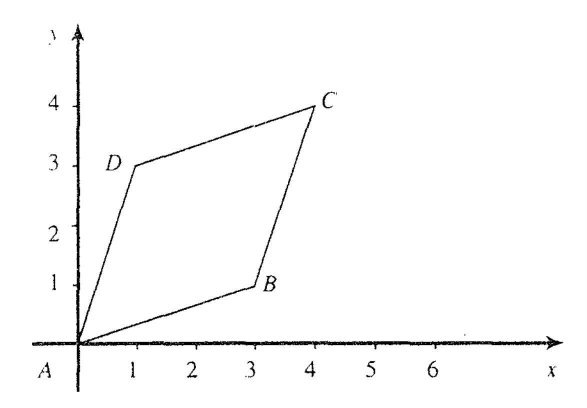


Figura 6. La región del ejemplo 4

Es claro que esta región no puede verse como una del tipo (I) o del tipo (II), pues si limitamos por izquierda y por derecha con rectas del tipo x = constante, que en este caso serían $x = 0$ y $x = 4$ (para tratar de verla como una región del tipo (I)), por arriba y por debajo nos quedarían dos funciones de x distintas en cada caso limitando la región (por debajo $\phi_1(x)$ y $\phi_2(x)$, y por arriba $\phi_3(x)$ y $\phi_4(x)$). Lo mismo ocurre si limitamos por arriba y por debajo por rectas ($y = 0$, $y = 4$). Entonces, ésta es una región del tipo (III): debemos dividirla para verla como la unión de varias regiones del tipo (I) o (II). De hecho es fácil ver que considerando las rectas $x = 1$ y $x = 3$ (justamente donde se juntan las dos funciones distintas que limitan la región por arriba y por debajo), nos queda la región dividida en 3 subregiones del tipo (I), a saber:

$$D_1 = \{(x, y) \in \mathbb{R}^2 | 0 \le x \le 1, x/3 \le y \le 3x\}$$
$$D_2 = \{(x, y) \in \mathbb{R}^2 | 1 \le x \le 3, x/3 \le y \le (x + 8)/3\}$$
$$D_3 = \{(x, y) \in \mathbb{R}^2 | 3 \le x \le 4, 3x - 8 \le y \le (x + 8)/3\}$$

Nótese, sin embargo, que la región se hubiera podido partir horizontalmente con las rectas $y = 1$ y $y = 3$, y tenerla así dividida en 3 subregiones del tipo (II), a saber

$$R_1 = \{(x, y) \in \mathbb{R}^2 | 0 \le y \le 1, y/3 \le x \le 3y\}$$
$$R_2 = \{(x, y) \in \mathbb{R}^2 | 1 \le y \le 3, y/3 \le x \le (y + 8)/3\}$$
$$R_3 = \{(x, y) \in \mathbb{R}^2 | 3 \le y \le 4, 3y - 8 \le x \le (y + 8)/3\}$$

Estudiaremos ahora el problema de la integración de funciones $f(x, y)$ sobre regiones más generales, de las que se han estudiado anteriormente. Sea $f: R \to \mathbb{R}$ una función definida en una región R del plano $\mathbb{R}^2$ como las estudiadas en esta sección. Supongamos que f es continua en R (o bien,

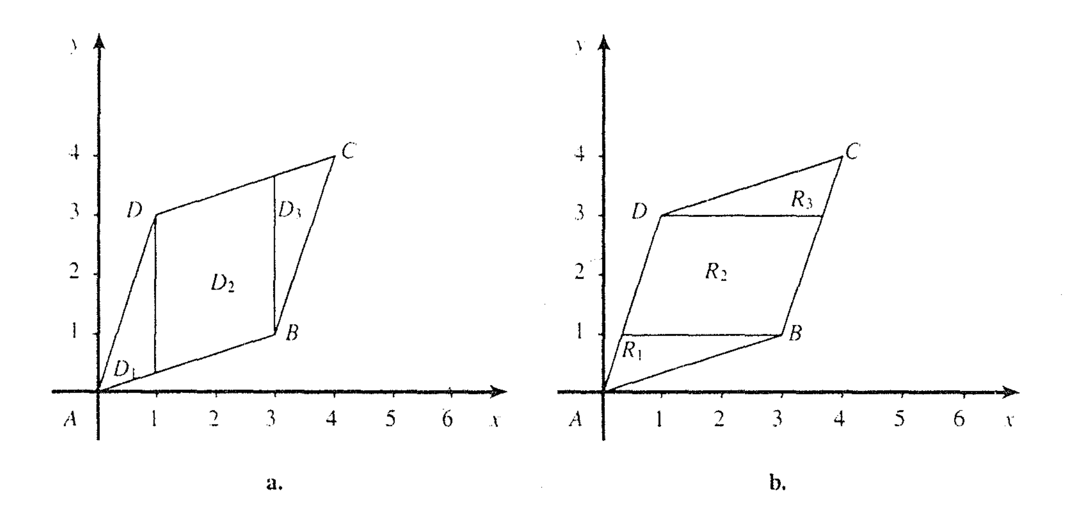


Figura 7. La región del ejemplo 4 vista como la unión de tres regiones **a.** del tipo (I), **b.** del tipo (II) ■

supongamos que al menos es continua en el interior de R y acotada en R). Siendo R una región acotada en el plano (¡por definición!), existe un rectángulo Q, que la contiene. Considere la función $\tilde{f}: Q \to \mathbb{R}$ definida como

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in R \\ 0 & \text{si } (x, y) \in Q \backslash R \end{cases}$$

Nótese que esta función puede llegar a ser discontinua en la frontera de la región R, que está constituida por una unión de gráficas de funciones continuas (funciones del tipo $y = \phi(x)$ o bien $x = \psi(y)$). Este conjunto de discontinuidades tiene medida cero (ver teorema 6.2.3 y el ejemplo 4 en la sección anterior). Entonces, por el teorema 6.2.4 de la sección anterior, la función $\tilde{f}(x, y)$ es integrable en Q. Se define así la integral doble de la función $f(x, y)$ sobre la región R, denotada por $\iint_R f(x, y)\,dx dy$, como

$$\iint_R f(x, y)\,dx\,dy = \iint_Q \tilde{f}(x, y)\,dx\,dy$$

Además, si la región R es del tipo I, es decir, si

$$R = \{(x, y)|a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$

donde $\phi_1, \phi_2: [a, b] \to$ son funciones continuas en el intervalo $[a, b]$, entonces la integral doble de $f(x, y)$ sobre R se puede calcular como

$$\iint_R f(x, y)\,dx\,dy = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\,dy\right)dx$$

mas si la región R es del tipo II, es decir, si

$$R = \{(x, y) | \psi_1(y) \leq x \leq \psi_2(y), c \leq y \leq d\}$$

donde $\psi_1, \psi_2: [c, d] \to \mathbb{R}$ son funciones continuas en $[c, d]$, entonces la integral doble de $f(x, y)$ sobre R se puede calcular como

$$\iint_R f(x, y)\, dx\, dy = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx \right) dy$$

(cada una de las integrales que aparecen en las expresiones anteriores, evaluadas de constante a constante, y de función de x —o de y— a función de x —o de y—, se les llama *integrales iteradas*). Vamos a dar un argumento detallado que muestre la primera de estas dos últimas afirmaciones. La segunda se demuestra de manera completamente análoga. Usaremos el teorema 6.2.1 de la sección anterior. Digamos en principio que el rectángulo Q donde está contenida la región R considerada (que es del tipo I) se puede tomar como $Q = [a, b] \times [\gamma, \delta]$, donde

$$\delta > \text{máx}(\phi_2(x), a \leq x \leq b), \quad \gamma < \text{mín}(\phi_1(x), a \leq x \leq b)$$

En esquema

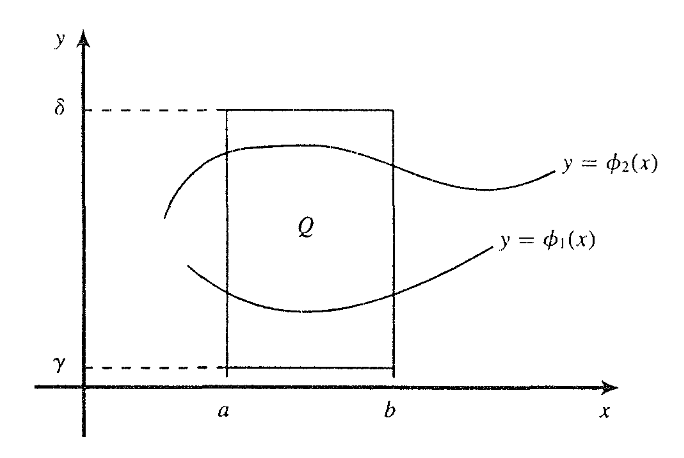


Figura 8. El rectángulo $Q = [a, b] \times [\gamma, \delta]$ que contiene a la región R

Para cada $x \in [a, b]$ fija, la función $\tilde{f}(x, y)$, $\gamma \leq y \leq \delta$ tiene cuando mucho 2 discontinuidades, a saber, en $y = \phi_1(x)$, $y = \phi_2(x)$, pues estamos suponiendo que la función $f(x, y)$ es continua en el interior de R. En tal caso, esta función $\tilde{f}(x, y)$ (vista como función de la variable y) es integrable en el intervalo $[\gamma, \delta]$, es decir, existe

$$h(x) = \int_\gamma^\delta \tilde{f}(x, y)\, dy$$

Además, obsérvese que

$$\begin{aligned} h(x) &= \int_{\gamma}^{\delta} \tilde{f}(x, y)\, dy \\ &= \int_{\gamma}^{\phi_1(x)} \tilde{f}(x, y)\, dy + \int_{\phi_1(x)}^{\phi_2(x)} \tilde{f}(x, y)\, dy + \int_{\phi_2(x)}^{\delta} \tilde{f}(x, y)\, dy \\ &= \int_{\gamma}^{\phi_1(x)} 0\, dy + \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy + \int_{\phi_2(x)}^{\delta} 0\, dy \\ &= \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy \end{aligned}$$

Tenemos así cumplida la primera parte de la hipótesis del teorema 6.2.1 (para cada $x \in [a, b]$, existe $h(x)$). Debemos verificar ahora que la función $h\colon [a, b] \to \mathbb{R}$ es integrable en $[a, b]$. Para esto, vamos a demostrar que esta función es de hecho *continua* en $[a, b]$ (y por lo tanto será integrable). Sin entrar mucho en detalles, exponemos a continuación los "pasos" del argumento que prueba la continuidad de $h(x)$:

$$\begin{aligned} |h(x+\alpha) - h(x)| &= \left| \int_{\phi_1(x+\alpha)}^{\phi_2(x+\alpha)} f(x+\alpha, y)\, dy - \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy \right| \\ &= \left| \int_{\phi_1(x+\alpha)}^{\phi_1(x)} f(x+\alpha, y)\, dy + \int_{\phi_1(x)}^{\phi_2(x)} f(x+\alpha, y)\, dy \right. \\ &\quad \left. + \int_{\phi_2(x)}^{\phi_2(x+\alpha)} f(x+\alpha, y)\, dy - \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy \right| \\ &\le \left| \int_{\phi_1(x+\alpha)}^{\phi_1(x)} f(x+\alpha, y)\, dy \right| + \left| \int_{\phi_2(x)}^{\phi_2(x+\alpha)} f(x+\alpha, y)\, dy \right| \\ &\quad + \int_{\phi_1(x)}^{\phi_2(x)} |f(x+\alpha, y) - f(x, y)|\, dy \\ &= |f(x+\alpha, \xi)||\phi_1(x) - \phi_1(x+\alpha)| + |f(x+\alpha, \zeta)||\phi_2(x+\alpha) - \phi_2(x)| \\ &\quad + \int_{\phi_1(x)}^{\phi_2(x)} |f(x+\alpha, y) - f(x, y)|\, dy \end{aligned}$$

en donde en el último paso hemos aplicado el teorema del valor medio, siendo ξ (ζ respectivamente) un número entre $\phi_1(x)$ y $\phi_1(x+\alpha)$ (entre $\phi_2(x)$ y $\phi_2(x+\alpha)$, respectivamente). Nótese que, de la última expresión, si α tiende a cero, los dos primeros términos (sus segundos factores) tienden a cero por la continuidad de las funciones ϕ_1 y ϕ_2. El tercer término también tiende a cero por la continuidad de la función $f(x, y)$ (en particular por la continuidad de esta función en su primera variable). Así pues, hemos probado que cuando α tiende a cero, la expresión $|h(x+\alpha) - h(x)|$ tiende a cero. Esto significa que $h(x)$ es continua.

Con esto hemos terminado de verificar las hipótesis correspondientes del teorema 6.2.1. Así pues, concluimos que la función $\tilde{f}(x, y)$ es integrable en el rectángulo Q y que su integral se calcula como

$$\iint_Q \tilde{f}(x, y)\, dx\, dy = \int_a^b h(x)\, dx$$

Es decir, se tiene que

$$\iint_R f(x, y)\, dx\, dy = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy dx$$

como queríamos probar.

NOTA 1: En las expresiones

$$\iint_R f(x, y)\, dx\, dy = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy \right) dx$$

y

$$\iint_R f(x, y)\, dx\, dy = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx \right) dy$$

podemos hacer caso omiso de los paréntesis, pues queda claro, por ejemplo en la primera de ellas, que la primera integral la tenemos que hacer *respecto de la variable* y, ya que los límites de integración son funciones de x y así se asegura que el resultado de esa primera integración será una función de x que tendremos que integrar finalmente (respecto de x) entre a y b. Así, podemos escribir las expresiones anteriores como

$$\iint_R f(x, y)\, dx\, dy = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy dx = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dx\, dy$$

y

$$\iint_R f(x, y)\, dx\, dy = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y)\, dx\, dy$$

respectivamente.

NOTA 2: En la discusión anterior se pidió que la función f fuera continua en la región R (o al menos continua en el interior de R y acotada en R). Esto se aplicó, por ejemplo, para ver que la función $\tilde{f}(x, y)$ tenía sólo dos posibles discontinuidades y así, asegurar que la función $h(x) = \int_\gamma^\delta \tilde{f}(x, y)\, dy$ fuera integrable. En realidad el resultado establecido con f continua en R sigue siendo válido si permitimos que esta función tenga discontinuidades en un subconjunto $\tilde{R}$ de R de medida cero. Es decir, vale un resultado análogo al teorema 6.2.4 del apéndice de la sección anterior (en el que se establece que "si $f: Q \to \mathbb{R}$ es una función definida y acotada en el rectángulo Q de $\mathbb{R}^2$, de modo que el conjunto de discontinuidades de f en Q tiene medida cero; entonces f es integrable en Q"), pero ahora para regiones más generales: si $f: R \to \mathbb{R}$ es una función definida y acotada en la región R de $\mathbb{R}^2$, de modo que el conjunto de discontinuidades de f en R tiene medida cero; entonces f es integrable en R. Es un hecho, si bien no lo usaremos en este capítulo pues las funciones que normalmente nos aparecerán en "la práctica" son continuas, que dejamos asentado, pues en él se dan las condiciones más generales para que una función acotada en una región R (no necesariamente un rectángulo) sea integrable. Tal resultado llega a ser necesario para justificar algunas discusiones teóricas que aparecerán más tarde (por ejemplo, en la sección 1 del capítulo 9 haremos referencia a él).

Ejemplo 5. Consideremos una función continua $f(x, y)$ en la región R estudiada en el ejemplo 1

$$R = \{(x, y)| -2 \leq x \leq 2, x^2 - 4 \leq y \leq 4 - x^2\}$$

Entonces la integral doble de $f(x, y)$ en R se calcularía como

$$\iint_R f(x, y)\, dy dx = \int_{-2}^{2} \int_{x^2-4}^{4-x^2} f(x, y)\, dy dx$$

en donde se hace primero la integral respecto de y (evaluándose entre $x^2 - 4$ y $4 - x^2$) y posteriormente respecto de x (evaluándose entre -2 y 2). ■

Ejemplo 6. Sea $f(x, y)$ una función continua en la región descrita por el trapecio cuyos vértices son $A = (1, 1)$, $B = (4, 1)$, $C = (2, 4)$ y $D = (3, 4)$ (ver ejemplo 2). Esta región es del tipo II y queda descrita como

$$R = \{(x, y)| 1 \leq y \leq 4, (y + 2)/3 \leq x \leq (13 - y)/3\}$$

La integral doble de $f(x, y)$ en R se calcula entonces como

$$\iint_R f(x, y)\, dy dx = \int_{1}^{4} \int_{(y+2)/3}^{(13-y)/3} f(x, y)\, dx\, dy$$

donde primero se hace la integral respecto de x (evaluándose entre $(13 - y)/3$ y $(y + 2)/3$) y luego respecto de y (evaluándose entre 1 y 4). ■

Ejemplo 7. Sea $f(x, y)$ una función continua definida en el círculo $x^2 + y^2 \leq r^2$. Esta región la podemos ver como del tipo I o del tipo II. En efecto, la región puede quedar descrita como

$$R = \{(x, y)| -r \leq x \leq r, -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}\}$$

en cuyo caso la integral doble de $f(x, y)$ en R se calcularía como

$$\iint_R f(x, y)\, dx\, dy = \int_{-r}^{r} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} f(x, y)\, dy\, dx$$

bien puede quedar descrita como

$$R = \{(x, y)| -r \leq y \leq r, -\sqrt{r^2 - y^2} \leq x \leq \sqrt{r^2 - x^2}\}$$

en cuyo caso la integral doble de $f(x, y)$ en R se calcularía como

$$\iint_R f(x, y)\, dx\, dy = \int_{-r}^{r} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} f(x, y)\, dx\, dy$$ ■

Las propiedades establecidas en el teorema 6.1.1 para las integrales dobles de funciones escalonadas definidas en rectángulos, son válidas en general para las funciones continuas que hemos considerado en esta sección, definidas en regiones más generales. Enunciamos el teorema que establece estas propiedades

Teorema 6.3.1 Sean $f: R \subset \mathbb{R}^2 \to \mathbb{R}$, $g: R \subset \mathbb{R}^2 \to \mathbb{R}$ dos funciones continuas definidas en la región R.

a. Si α y β son dos números reales cualesquiera, entonces

$$\int_R\int(\alpha f(x, y) + \beta g(x, y))\,dx\,dy = \alpha\int_R\int f(x, y)\,dx\,dy + \beta\int_R\int g(x, y)\,dx\,dy$$

b. Si $f(x, y) \geq g(x, y)$ para toda $(x, y) \in R$, entonces

$$\iint\limits_R f(x, y)\,dx\,dy \geq \int_R\int g(x, y)\,dx\,dy$$

c. Si la región R está dividida en dos subregiones R_1 y R_2 (es decir, si se tiene $R = R_1 \cup R_2$), entonces

$$\iint\limits_R f(x, y)\,dx\,dy = \iint\limits_{R_1} f(x, y)\,dx\,dy + \iint\limits_{R_2} f(x, y)\,dx\,dy$$ ■

Ejemplo 8. Sea $f: [a, b] \to \mathbb{R}$ una función continua en el intervalo $[a, b]$. Consideremos la función $F: Q \to \mathbb{R}$ definida en el rectángulo $Q = [a, b]^2$ de $\mathbb{R}^2$ dada por $F(x, y) = f(x) - f(y)$. De hecho se tiene que $(F(x, y))^2 \geq 0$ para toda $(x, y) \in Q$, de modo que la propiedad **c** del teorema anterior nos dice que

$$\iint\limits_Q (F(x, y))^2 dx\,dy \geq 0$$

Más aún, se tiene que $(F(x, y))^2 > 0$ para toda $(x, y) \in Q$ y $(F(x, y))^2 = 0$ si y sólo si $F(x, y) = 0$ para toda $(x, y) \in Q$. Veamos cómo de esta simple observación se puede obtener una desigualdad interesante para la función $f(x)$. Escribiendo explícitamente la función $F(x, y)$ y haciendo uso de lo estudiado en esta sección acerca de cómo se calculan las integrales dobles (o bien del teorema 6.2.1), así como de la propiedad **a** del teorema anterior, nos queda

$$\begin{aligned}
0 &\leq \iint\limits_Q (F(x, y))^2 dx\,dy = \int_a^b\int_a^b (f(x) - f(y))^2\,dx\,dy \\
&= \int_a^b\int_a^b (f^2(x) - 2f(x)f(y) + f^2(y))\,dx\,dy \\
&= \int_a^b\int_a^b f^2(x)\,dx\,dy - 2\int_a^b\int_a^b f(x)f(y)\,dx\,dy + \int_a^b\int_a^b f^2(y)\,dx\,dy \\
&= \int_a^b f^2(x)\,dx\int_a^b dy - 2\int_a^b f(x)\,dx\int_a^b f(y)\,dy + \int_a^b dx\int_a^b f^2(y)\,dy \\
&= (b - a)\int_a^b f^2(x)\,dx - 2\left(\int_a^b f(x)dx\right)^2 + (b - a)\int_a^b f^2(x)\,dx \\
&= 2(b - a)\int_a^b f^2(x)\,dx - 2\left(\int_a^b f(x)\,dx\right)^2
\end{aligned}$$

de donde se obtiene entonces la interesante desigualdad para la función $f(x)$

$$\left(\int_a^b f(x)\,dx\right)^2 \leq (b-a)\int_a^b f^2(x)\,dx$$

de la cual podemos asegurar además que la igualdad se da si y sólo si la función $f(x)$ es constante en el intervalo $[a, b]$ (pues como se vió, la igualdad $\iint_Q (F(x, y))^2\,dx\,dy = 0$ se da si y sólo si $F(x, y) = f(x) - f(y) = 0$ para toda $(x, y) \in Q$, es decir, si y sólo si $f(x) = f(y)$ para toda x, $y \in [a, b]$, lo cual significa que la función $f(x)$ es constante en el intervalo $[a, b]$). ■

Ejemplo 9. Sea $f(x, y)$ una función continua en la región R que consiste en el triángulo cuyos vértices son $A = (0, 0)$, $B = (4, 0)$ y $C = (2, 2)$.

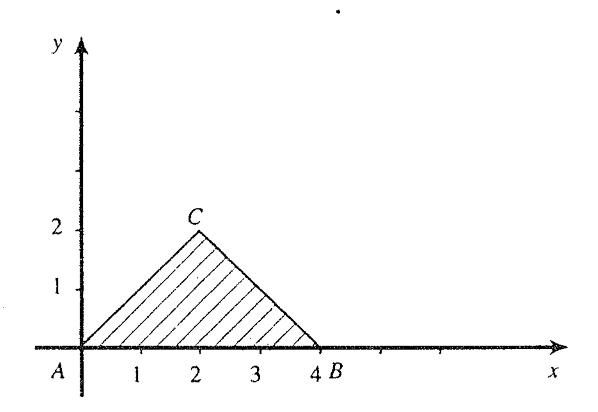


Figura 9. Región R para el ejemplo 9.

De manera natural esta región se ve como del tipo II: limitada por debajo por la recta $y = 0$, por encima por la recta $y = 2$, por la izquierda por la recta $x = \psi_1(y) = y$, y por la derecha por la recta $x = \psi_2(y) = 4 - y$, de tal manera que la integral de $f(x, y)$ sobre R se calcularía como

$$\iint_R f(x, y)\,dx\,dy = \int_0^2 \int_y^{4-y} f(x, y)\,dx\,dy$$

Se entiende que en esta integral doble primero se hace la integración respecto de x y luego respecto de y. Si quisiéramos *invertir el orden de integración*, la región R deberá verse como una región del tipo I, o bien, como una unión de regiones de este tipo. En este caso no es posible ver nuestra región como del tipo I, pues su "techo" lo constituyen dos funciones de x diferentes (la función $y = \phi_1(x) = x$ en $0 \leq x \leq 2$, y $y = \phi_2(x) = 4 - x$ en $2 \leq x \leq 4$). Sin embargo, dividiendo nuestra región R por la recta $x = 2$, la podemos considerar como la unión de dos subregiones del tipo I, a saber

$$R_1 = \{(x, y) | 0 \leq x \leq 2, 0 \leq y \leq x\}$$
$$R_2 = \{(x, y) | 2 \leq x \leq 4, 0 \leq y \leq 4 - x\}$$

En tal caso, haciendo uso de la propiedad c) del teorema 6.3.1, podemos escribir

$$\begin{aligned}\iint_R f(x, y)\,dx\,dy &= \iint_{R_1} f(x, y)\,dx\,dy + \iint_{R_2} f(x, y)\,dx\,dy \\ &= \int_0^2 \int_0^x f(x, y)\,dydx + \int_2^4 \int_x^{4-x} f(x, y)\,dydx\end{aligned}$$ ■

Ejemplo 10. Retomemos el ejemplo 6. Queremos invertir el orden de integración en la integral doble de la función continua $f(x, y)$ sobre el trapecio de vértices $A = (1, 1)$, $B = (4, 1)$, $C = (2, 4)$ y $D = (3, 4)$. Para lograrlo, debemos de partir esta región en 3 subregiones R_1, R_2 y R_3 del tipo I con las rectas verticales $x = 2$ y $x = 3$. Estas subregiones son entonces

$$\begin{aligned}R_1 &= \{(x, y)|1 \leq x \leq 2, 1 \leq y \leq 3x - 2\} \\ R_2 &= \{(x, y)|2 \leq x \leq 3, 1 \leq y \leq 4\} \\ R_3 &= \{(x, y)|3 \leq x \leq 4, 1 \leq y \leq 13 - 3x\}\end{aligned}$$

de modo que la integral de $f(x, y)$ sobre la región $R = R_1 \cup R_2 \cup R_3$ se calcularía como (usando la propiedad **c.** del teorema 6.3.1)

$$\begin{aligned}\iint_R f(x, y)\,dx\,dy &= \iint_{R_1} f(x, y)\,dx\,dy + \iint_{R_2} f(x, y)\,dx\,dy + \iint_{R_3} f(x, y)\,dx\,dy \\ &= \int_1^2 \int_1^{3x-2} f(x, y)\,dydx + \int_2^3 \int_1^4 f(x, y)\,dydx + \int_3^4 \int_1^{13-x} f(x, y)\,dydx\end{aligned}$$ ■

Ejemplo 11. Supongamos que se quiere calcular la integral

$$\int_0^1 \int_y^{2-\sqrt{y}} f(x, y)\,dx\,dy$$

Por la manera como están escritos los límites de integración, es claro que se trata de la integral doble de la función $f(x, y)$ en una región R del tipo II, a saber

$$R = \{(x, y)|0 \leq y \leq 1, y \leq x \leq 2 - \sqrt{y}\}$$

limitada por debajo por la recta $y = 0$ (el eje de las x), por arriba por la recta $y = 1$, por la izquierda por la recta $x = \psi_1(y) = y$, y por la derecha por la parábola $x = \psi_2(y) = 2 - \sqrt{y}$. También es claro que primero se debe integrar respecto de la variable x y luego respecto de y. Esta región se ve como (figura 10)

Supongamos que se quiere invertir el orden de integración. Es decir, se quiere integrar primero respecto de y y luego respecto de x. Para esto habrá que ver la región de integración R como una región del tipo I, o bien, como una unión de ellas. Es claro que en este caso habrá que partir la región R con la recta $x = 1$, dejando así dos regiones R_1 y R_2 cuya unión es R. Estas son

$$\begin{aligned}R_1 &= \{(x, y)|0 \leq x \leq 1, 0 \leq y \leq x\} \\ R_2 &= \{(x, y)|1 \leq x \leq 2, 0 \leq y \leq (x - 2)^2\}\end{aligned}$$

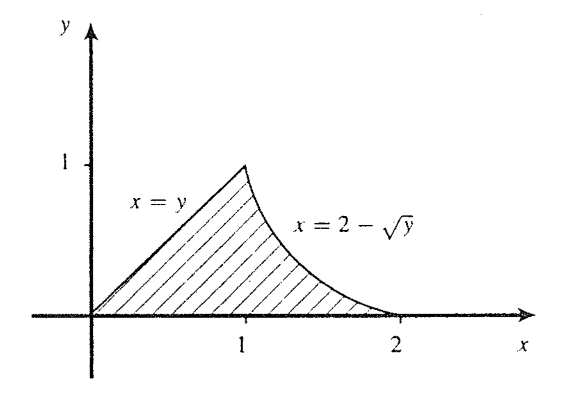


Figura 10. La región de integración para el ejemplo 11.

Se tiene entonces que

$$\int_0^1 \int_y^{2-\sqrt{y}} f(x, y)\,dx\,dy = \int_0^1 \int_0^x f(x, y)\,dydx + \int_1^2 \int_0^{(x-2)^2} f(x, y)\,dydx \qquad \blacksquare$$

El problema de dejar establecidos los conceptos teóricos sobre integrales dobles de funciones en regiones del plano $\mathbb{R}^2$, ha quedado, con esta sección, cubierto lo más decorosamente posible. Lo importante en este momento es que se desarrollen habilidades operativas para poder *calcular efectivamente* integrales dobles. Al calcular una integral doble de una función $f(x, y)$ en una región R debemos ser capaces de: (1) saber describir bien la región R y poner los límites correspondientes a las integrales simples por medio de las cuales se calculará la integral doble, y, (2) saber calcular en forma efectiva las integrales involucradas. El punto (2) se puede cubrir recordando las técnicas de integración de nuestro primer curso de cálculo. Para adquirir destreza en el punto (1), se tiene que trabajar con muchos ejemplos concretos. El resto de esta sección lo dedicaremos a resolver algunos de ellos.

Ejemplo 12. Se quiere calcular la integral de la función $f(x, y) = \operatorname{sen}(x + y)$ en la región acotada por las rectas $x = 0$, $y = 3\pi$, $y = x$.

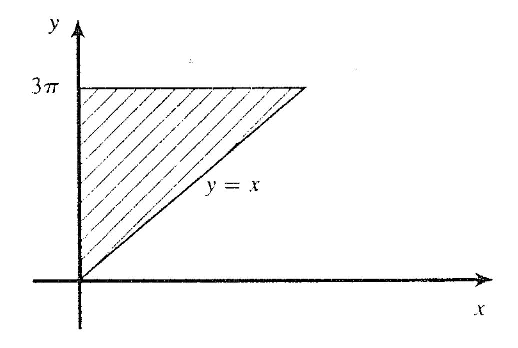


Figura 11. La región de integración para el ejemplo 12.

La región puede ser considerada del tipo I o del tipo II. Si se la ve como una región del tipo II queda

$$R = \{(x, y)|0 \leq y \leq 3\pi, 0 \leq x \leq y\}$$

de modo que la integral se calcularía como

$$\iint_R f(x, y)\,dx\,dy = \int_0^{3\pi}\int_0^y \operatorname{sen}(x+y)\,dx\,dy = \int_0^{3\pi} -\cos(x+y)\Big|_0^y dy$$

$$= -\int_0^{3\pi}(\cos(2y) - \cos y)\,dy = \operatorname{sen} y - \frac{1}{2}\operatorname{sen}(2y)\Big|_0^{3\pi} = 0 \qquad \blacksquare$$

Ejemplo 13. Calculemos la integral $\iint_R f(x, y)\,dx\,dy$ de la función

$$f(x, y) = \sqrt{r^2 - x^2 - y^2}$$

donde la región R es $x^2 + y^2 \leq r^2$. Obsérvese que esta región es el interior del círculo $x^2 + y^2 = r^2$ (incluyendo su frontera), y se puede ver como una región del tipo I (ver ejemplo 7), a saber,

$$R = \{(x, y)| -r \leq x \leq r, -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}\}$$

Entonces

$$\iint_R f(x, y)\,dx\,dy = \int_{-r}^{r}\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \sqrt{r^2 - x^2 - y^2}\,dydx$$

Es claro que la primera integral se debe hacer respecto de y, la cual es una integral de la forma

$$\int \sqrt{a^2 - y^2}\,dy$$

con la "constante" $a^2 = r^2 - x^2$ (que, aunque depende de x, al integrar respecto de y se considera como constante). Esta integral se puede calcular haciendo $y = a \operatorname{sen} t$. En tal caso

$$\int \sqrt{a^2 - y^2}\,dy = \int (a\cos t)(a\cos t)\,dt = a^2 \int \cos^2 t\,dt$$

$$= a^2 \int \frac{1}{2}(1 + \cos 2t)\,dt = \frac{a^2}{2}(t + \frac{1}{2}\operatorname{sen} 2t)$$

$$= \frac{a^2}{2}(t + \operatorname{sen} t \cos t) = \frac{a^2}{2}\left(\operatorname{arcsen}\frac{y}{a} + \frac{y}{a}\frac{\sqrt{a^2 - y^2}}{a}\right)$$

$$= \frac{a^2}{2}\operatorname{arcsen}\frac{y}{a} + \frac{y}{2}\sqrt{a^2 - y^2}$$

Entonces

$$\begin{aligned}
\iint\limits_R f(x, y)\,dx\,dy &= \int_{-r}^{r} \left.\frac{r^2 - x^2}{2} \operatorname{arcsen} \frac{y}{\sqrt{r^2 - x^2}} + \frac{y}{2}\sqrt{r^2 - x^2 - y^2}\right|_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} dx \\
&= \int_{-r}^{r} \frac{r^2 - x^2}{2} (\operatorname{arcsen} 1 - \operatorname{arcsen}(-1))\,dx = \frac{\pi}{2}\int_{-r}^{r} (r^2 - x^2)\,dx \\
&= \frac{\pi}{2}\left(r^2 x - \frac{1}{3}x^3\right)_{-r}^{r} = \frac{2}{3}\pi r^3
\end{aligned}$$

Quisiéramos llamar la atención al segnificado geométrico del resultado obtenido. Obsérvese que la función que integramos $f(x, y)$ tiene por gráfica el hemisferio superior de la esfera $x^2 + y^2 + z^2 = r^2$ y la región de integración es justamente la proyección de este hemisferio sobre el plano xy (es decir, la región $x^2 + y^2 \leq r^2$). Siendo esta función $f(x, y)$ no negativa y continua en R, el resultado obtenido del cálculo de su integral doble en R se puede interpretar como el volumen bajo la superficie de $f(x, y)$. Así pues, lo que se calculó fue el volumen de media esfera de radio r (el volumen total de esta esfera es, como sabemos, $V = \frac{4}{3}\pi r^3$, y el resultado que obtuvimos es justamente $\frac{1}{2}V$). ■

Ejemplo 14. Calculemos la integral $\iint_R (x + y)\,dx\,dy$ donde R es la cuarta parte, situada en el primer cuadrante, del anillo circular limitado por los círculos $x^2 + y^2 = 1$, $x^2 + y^2 = 4$.

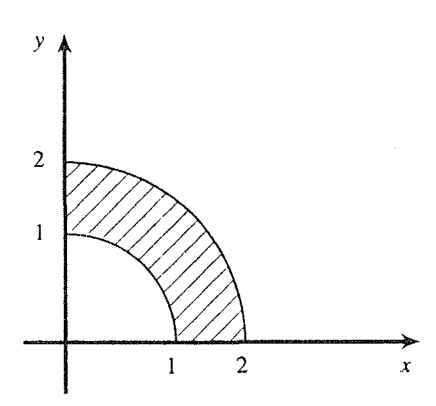


Figura 12. La región de integración para el ejemplo 14.

Esta región no se puede ver globalmente como una región del tipo I o del tipo II. Hay que partirla en dos subregiones, ya sea con la recta $x = 1$ (en cuyo caso las subregiones resultantes son del tipo I), o con la recta $y = 1$ (en cuyo caso las subregiones resultantes son del tipo II). Tomemos la primera de estas opciones. Entonces $R = R_1 \cup R_2$, donde

$$R_1 = \{(x, y)|0 \leq x \leq 1, \sqrt{1 - x^2} \leq y \leq \sqrt{4 - x^2}\}$$
$$R_2 = \{(x, y)|1 \leq x \leq 2, 0 \leq y \leq \sqrt{4 - x^2}\}$$

Se tiene entonces

$$\iint_R (x+y)\,dx\,dy = \iint_{R_1} (x+y)\,dx\,dy + \iint_{R_2} (x+y)\,dx\,dy$$

$$= \int_0^1 \int_{\sqrt{1-x^2}}^{\sqrt{4-x^2}} (x+y)\,dy\,dx + \int_1^2 \int_0^{\sqrt{4-x^2}} (x+y)\,dy\,dx$$

$$= \int_0^1 x\left(\sqrt{4-x^2} - \sqrt{1-x^2}\right) dx + \int_0^1 \frac{1}{2}\left(4 - x^2 - (1-x^2)\right) dx$$

$$+ \int_1^2 x\sqrt{4-x^2}\,dx + \int_1^2 \frac{1}{2}(4-x^2)\,dx$$

$$= \left[-\frac{1}{3}(4-x^2)^{3/2} + \frac{1}{3}(1-x^2)^{3/2} + \frac{3}{2}x\right]_0^1 + \left[-\frac{1}{3}(4-x^2)^{3/2} + 2x - \frac{1}{6}x^3\right]_1^2$$

$$= \frac{14}{3}$$ ■

Ejemplo 15. Se quiere calcular la integral de la función $f(x, y) = x + y + 1$ sobre la región R limitada por las rectas $y - x = 1$, $y - x = -1$, $y + x = 1$, $y + x = 2$. Es fácil encontrar, resolviendo simultáneamente por pares estas ecuaciones, los puntos donde se intersectan estas rectas. Estos son $A = (0, 1)$, $B = (1, 0)$, $C = (3/2, 1/2)$, $D = (1/2, 3/2)$. La región R se ve entonces como

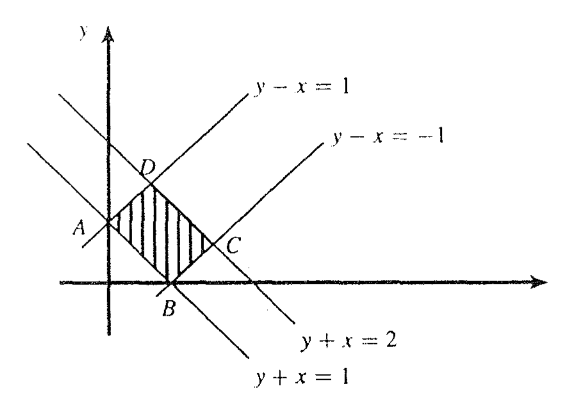


Figura 13. La región de integración para el ejemplo 15.

Es patente que R no es (globalmente) una región del tipo I ni del tipo II. Habrá que dividirla por las rectas $x = 1/2$, $x = 1$ para verla como una unión de 3 subregiones del tipo I, o bien, por las rectas $y = 1/2$, $y = 1$ para verla como una unión de 3 subregiones del tipo II. Hagámoslo de la primera de estas maneras. Entonces como $R = R_1 \cup R_2 \cup R_3$, donde

$$R_1 = \{(x, y) | 0 \le x \le 1/2,\ 1 - x \le y \le 1 + x\}$$
$$R_2 = \{(x, y) | 1/2 \le x \le 1,\ 1 - x \le y \le 2 - x\}$$
$$R_3 = \{(x, y) | 1 \le x \le 3/2,\ x - 1 \le y \le 2 - x\}$$

La integral doble de $f(x, y)$ sobre R se calcula como

$$\begin{aligned}
\iint_R f(x, y)\,dx\,dy &= \iint_{R_1} f(x, y)\,dx\,dy + \iint_{R_2} f(x, y)\,dx\,dy + \iint_{R_3} f(x, y)\,dx\,dy \\
&= \int_0^{1/2}\int_{1-x}^{1+x} (x + y + 1)\,dydx + \int_{1/2}^{1}\int_{1-x}^{2-x} (x + y + 1)\,dydx \\
&\quad + \int_1^{3/2}\int_{x-1}^{2-x} (x + y + 1)\,dydx \\
&= \int_0^{1/2} (x + 1)\,dx \int_{1-x}^{1+x} dy + \int_0^{1/2} dx \int_{1-x}^{1+x} ydy \\
&\quad + \int_{1/2}^{1} (x + 1)\,dx \int_{1-x}^{2-x} dy + \int_{1/2}^{1} dx \int_{1-x}^{2-x} y\,dy \\
&\quad + \int_1^{3/2} (x + 1)\,dx \int_{x-1}^{2-x} dy + \int_1^{3/2} dx \int_{x-1}^{2-x} y\,dy \\
&= \int_0^{1/2} (x + 1)(2x)\,dx + \int_0^{1/2} 2x\,dx + \int_{1/2}^{1} (x + 1)\,dx + \frac{1}{2}\int_{1/2}^{1} (3 - 2x)\,dx \\
&\quad + \int_1^{3/2} (x + 1)(3 - 2x)\,dx + \frac{1}{2}\int_1^{3/2} (3 - 2x)\,dx \\
&= \left[\frac{2}{3}x^3 + 2x^2\right]_0^{1/2} + \left[\frac{5}{2}x\right]_{1/2}^{1} + \left[-\frac{2}{3}x^3 + \frac{9}{2}x\right]_1^{3/2} = \frac{5}{2}
\end{aligned}$$

■

En la próxima sección se expondrá un "artificio" sumamente poderoso que nos permitirá calcular algunas integrales dobles de un modo muy eficiente: estudiaremos la problemática relacionada con "cambios de variables en integrales dobles".

Ejercicios (Capítulo 6, Sección 3)

En los ejercicios 1—10 se da una región R en el plano. Escriba los límites de las integrales iteradas con las que se calcula la integral doble $\iint_R f(x, y)\,dx\,dy$, donde $f(x, y)$ es una función continua en la región R. En todos los casos, la integral doble se puede calcular con *una* integral iterada, viendo que R convenga, como una región del tipo (I) o del tipo (II). En caso de que R se pueda ver como una región de ambos tipos, escriba las dos expresiones correspondientes de las integrales iteradas.

1. R es la región triangular limitada por los ejes coordenados y la recta $y = 4 - 3x$.
2. R es la región triangular limitada por el eje x, la recta $x = 4$, y la recta $y = 4x$.
3. R es el la región limitada por el triángulo cuyos vértices son $A = (1, 1)$, $B = (5, 1)$, $C = (5, 6)$.
4. R es la región limitada por el triángulo cuyos vértices son $A = (2, 3)$, $B = (6, 3)$, $C = (4, 6)$.
5. $R = \{(x, y)|x^2 + 2y^2 \leq 1\}$

6. $R = \{(x, y)|x^2 \leq y \leq \sqrt{x}\}$

7. $R = \{(x, y)|(x-1)^2 + (y-2)^2 \leq 1\}$

8. $R = \{(x, y)|x^2 \leq y \leq x+2\}$

9. $R = \{(x, y)|y^2 - 1 \leq x \leq 1 - y^2\}$

10. $R = \{(x, y)|y^2 + 2 \leq x \leq y + 3\}$

En los ejercicios 11—15, se da una región R en el plano. Para calcular la integral doble $\iint_R f(x, y)\,dxdy$, se necesita expresar R como la unión de subregiones y aplicar la propiedad **c** del teorema 6.3.1. Escriba en cada caso la correspondiente suma de integrales iteradas, viendo a las subregiones de R como regiones: **a.** del tipo (I), **b.** del tipo (II).

11. R es la región encerrada por el cuadrado $|x| + |y| = 1$

12. R es la región encerrada por el rombo con vértices en $A = (0, 0)$, $B = (2, 3)$, $C = (0, 6)$, $D = (-2, 3)$.

13. R es la región encerrada por el cuadrilátero con vértices en $A = (1, 0)$, $B = (10, 1)$, $C = (7, 5)$, $D = (3, 5)$.

14. R es la región encerrada por el paralelogramo con vértices en $A = (0, 0)$, $B = (5, 1)$, $C = (7, 5)$, $D = (2, 4)$.

15. R es la región encerrada por el pentágono con vértices en $A = (0, 0)$, $B = (3, -3)$, $C = (6, 0)$, $D = (6, 4)$, $E = (0, 4)$.

En los ejercicios 16—20, se da una (o una suma de) integral(es) iterada(s) de la función $f(x, y)$ sobre la región R (o sobre subregiones de R, respectivamente). Haga un dibujo que muestre la región de integración R y escriba la expresión de la integral iterada correspondiente, si se intercambiara el orden de integración.

16. $\int_0^1 dx \int_{x^2}^{\sqrt{x}} f(x, y)\,dy$

17. $\int_{-2}^2 dx \int_0^{0.5(4-x^2)^{1/2}} f(x, y)\,dy$

18. $\int_0^1 dx \int_0^{(1-x^2)^{1/2}} f(x, y)\,dy$

19. $\int_0^1 dx \int_0^x f(x, y)dy + \int_1^{10} dx \int_0^{(10-x)/9} f(x, y)\,dy$

20. $\int_{-2}^0 dx \int_0^{(x+2)^2} f(x, y)dy + \int_0^2 dx \int_0^{(x-2)^2} f(x, y)\,dy$

En los ejercicios 21—30, calcule la integral doble indicada.

21. $\iint\limits_R (x + 3y)\,dx\,dy$, donde R es la región limitada por el eje x, la recta $y = 2x$, y la recta $y = -x + 4$.

22. $\iint_R x^2y^3 dx\,dy$, donde R es el la región limitada por el rectángulo de vértices en $A = (0,0)$, $B = (3,0)$, $C = (3,2)$, $D = (0,2)$.

23. $\iint_R \frac{1}{x+y}\,dx\,dy$, donde R es la región limitada por las rectas $y = 0$, $x = 1$, $y = x$.

24. $\iint_R \operatorname{sen}(x+y)\,dx\,dy$, donde R es la región limitada por las rectas $x = 0$, $y = x$, $y = 3\pi$.

25. $\iint_R \frac{x^2}{y}\,dx\,dy$, donde R es la región limitada por las parábolas $y = x^2$, $x = y^2$.

26. $\iint_R (x^2+3y^3)^2\,dx\,dy$, donde R es la región limitada por las rectas $x = 0$, $x = 1$, $y = 1$, $y = 0$.

27. $\iint_R (2x+y)^3 dx\,dy$, donde R es la región limitada por el triángulo cuyos vértices son $A = (1,1)$, $B = (4,0)$, $C = (3,5)$.

28. $\iint_R e^x e^{2y}\,dx\,dy$, donde R es el la región limitada por el cuadrado $|x| + |y| = 1$.

29. $\iint_R x^2ye^{x-y}\,dx\,dy$, donde R es la región limitada por el rectángulo $[0,3] \times [0,2]$.

30. $\iint_R x^2y\operatorname{sen}(x+y)\,dx\,dy$, donde R es la región limitada por las rectas $x = 0$, $x = y$, $y = 1$.

31. Calcule la integral doble de la función $f(x,y) = \sqrt{1 - \frac{x^2}{a^2} - \frac{y_2}{b^2}}$ sobre la región limitada por la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

32. Calcule la integral doble de la función $f(x,y) = \operatorname{sgn}(x-y)$, sobre el cuadrado $|x| + |y| \le 1$.

33. Calcule la integral doble de la función $f(x,y) = \operatorname{sgn}(x^2-y)$, sobre el rectángulo $[-1,1]\times[0,1]$.

34. Calcule la integral doble de la función $f(x,y) = |x+y|$, sobre el rectángulo $[-1,1]\times[-1,1]$.

6.4 Cambio de variable en integrales dobles

Una de las técnicas poderosas que conocemos de nuestro primer curso de cálculo para calcular integrales, es la técnica del cambio de variables. Si queremos integrar la función $f(x)$ en el intervalo $[a,b]$, se trata de *cambiar la variable x por otra variable t*, según la ecuación $x = g(t)$ (llamada fórmula de cambio de variable), e integrar entonces la función $f(g(t))$ en otro intervalo $[c,d]$. Se

entiende que el cambio debe ser tal que la nueva función a integrar $f(g(t))$ debe ser "más sencilla" que la función original. De modo más precisamente, se tiene el resultado siguiente:

Sea $f: [a, b] \to \mathbb{R}$ *una función continua y* $g: [c, d] \to \mathbb{R}$ *una función derivable con derivada* $g'(t)$ *continua (es decir, g es de clase* $\mathscr{C}^1$*) tal que* $g([c, d]) \subset [a, b]$. *Entonces*

$$\int_{\phi(c)}^{\phi(d)} f(x)\,dx = \int_c^d f(g(t))g'(t)\,dt$$

Esta es la llamada "fórmula de cambio de variable" (en la integral definida) y, como decíamos, es un instrumento muy poderoso en la evaluación de algunas integrales. El objetivo de esta sección es estudiar el resultado análogo al anterior para integrales dobles de funciones $f(x, y)$ sobre regiones R del plano $\mathbb{R}^2$.

La problemática en este caso será más complicada que en el caso de funciones de una sola variable, pues, por ejemplo, ahora se trata de *cambiar las dos variables de la función* $f(x, y)$. En términos generales, se trata de cambiar las variables x, y en la función $f(x, y)$ por dos nuevas variables, digamos u, v, según las *ecuaciones de transformación*

$$x = \phi(u, v), \qquad y = \psi(u, v)$$

de tal modo que la integral original de la función $f(x, y)$ que se quiere hacer sobre la región R del plano $\mathbb{R}^2$ (en el que las variables son x, y -lo llamaremos "plano xy"), ahora se convierta en la integral de la nueva función $F(u, v) = f(\phi(u, v), \psi(u, v))$ sobre alguna nueva región R' del plano $\mathbb{R}^2$ (en el que las variables son u, v —lo llamaremos "plano uv").

Para esta presentación optaremos por rescatar la notación y terminología usada con anterioridad en el capítulo 3 (sección 6), en el que se estudió la problemática relacionada con las funciones inversas de funciones del tipo $F: R' \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $F(u, v) = (\phi(u, v), \psi(u, v))$, definidas en una región R' de $\mathbb{R}^2$. Así, en lugar de pensar en *dos* ecuaciones de transformación de coordenadas $x = \phi(u, v)$, $y = \psi(u, v)$ para la función $f(x, y)$, pensaremos en *la función F de transformación de coordenadas*, $F: R' \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $F(u, v) = (x, y) = (\phi(u, v), \psi(u, v))$, que manda puntos (u, v) de la región R' del plano uv, en puntos $(x, y) = (\phi(u, v), \psi(u, v))$ de una nueva región R del plano xy, en donde entonces las funciones ϕ y ψ son funciones definidas en $R' \subseteq \mathbb{R}^2$ y tomando valores reales. Bajo esta perspectiva, cambiar las variables x, y de la función $f(x, y)$, significa (¡al igual que en el caso de una sola variable!) *hacer la composición* de $f(x, y)$ con la función $(x, y) = F(u, v)$, obteniendo así la función compuesta $f(F(u, v))$, que será (como veremos a continuación, parte de) la nueva función que debemos integrar. En esquema esto se ve como (figura 1)

Enunciemos ahora con precisión el resultado que establece las condiciones sobre las cuales se hace el cambio de variables en una integral doble.

Teorema 6.4.1 Sea $f: R \subset \mathbb{R}^2 \to \mathbb{R}$ una función continua de las variables x, y definida en la región $R \subset \mathbb{R}^2$. Sea $F: R' \subset \mathbb{R}^2 \to \mathbb{R}^2$, $F(u, v) = (x, y) = (\phi(u, v), \psi(u, v))$ una función que manda de manera inyectiva los puntos (u, v) de la región R', en los puntos (x, y) de la región R del plano xy. Suponga que F es de clase $\mathscr{C}^1$ y que la derivada $F'(u, v)$ es una matriz inversible para todo $(u, v) \in R'$ (equivale a que el determinante $\det F'(u, v)$ sea no nulo en todos los puntos (u, v) de R'). Entonces se tiene la fórmula de cambio de variables en integrales dobles

$$\iint_R f(x, y)\,dx\,dy = \iint_{R'} f(\phi(u, v), \psi(u, v))\left|\frac{\partial(\phi, \psi)}{\partial(u, v)}\right| du\,dv \qquad \blacksquare$$

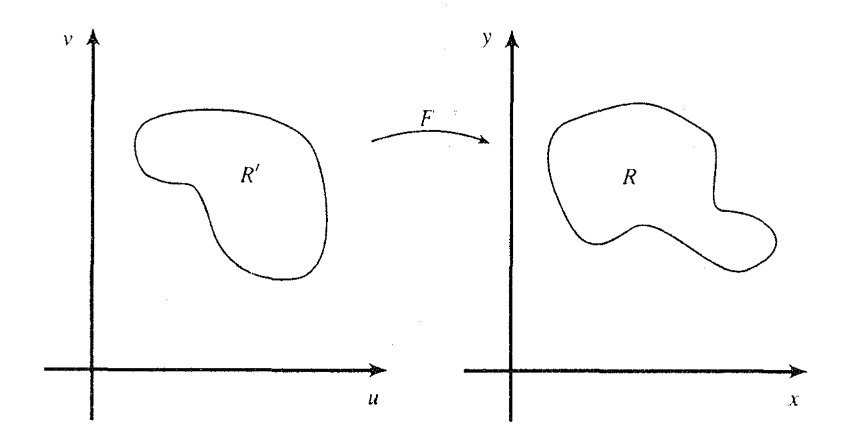


Figura 1. La función $F(u, v) = (x, y)$ que transforma la región R' en la región R

La demostración de este teorema no la presentaremos en esta sección. Hay muchas maneras de demostrar el resultado, que por ahora escapan al nivel alcanzado en el texto. Por una parte esto nos resulta incluso benéfico, pues en este momento lo que más importa es, después de entender bien lo que tal teorema establece, *desarrollar habilidades para aplicar adecuadamente la fórmula de cambio de variables*. Prometemos, sin embargo, que en el capítulo próximo, cuando se estudien las integrales de línea, presentar una demostración del teorema anterior usando el célebre Teorema de Green (uno de los principales resultados del siguiente capítulo. Ver apéndice (I) de la sección 8, capítulo 7).

De modo que en este momento haremos una serie de observaciones sobre el contenido del teorema anterior las cuales, junto con un argumento geométrico que nos hará plausible el resultado anterior (aunque sea en un caso particular), nos permitirán (así lo esperamos) entender su contenido. Luego, veremos algunos casos particulares importantes del teorema, y se resolverán un buen número de ejemplos concretos.

La primera observación que debemos hacer en la fórmula de cambio de variables del teorema anterior es la "sospechosa" aparición del determinante jacobiano

$$\frac{\partial(\phi, \psi)}{\partial(u, v)} = \det \begin{bmatrix} \dfrac{\partial \phi}{\partial u} & \dfrac{\partial \phi}{\partial v} \\ \dfrac{\partial \psi}{\partial u} & \dfrac{\partial \psi}{\partial v} \end{bmatrix}$$

o bien, de manera más sugestiva, escribiendo las funciones componentes de F como $x = x(u, v)$, $y = y(u, v)$, de modo que $F(u, v) = (x, y) = (x(u, v), y(u, v))$, el jacobiano se vería como

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix}$$

(así escribiremos este jacobiano de aquí en adelante). En realidad, si comparamos con el resultado ya conocido del cambio de variable en una integral simple, vemos que este jacobiano está jugando el

papel que desempeña la derivada $g'(t)$ de la ecuación de transformación $x = g(t)$. Recuerde que de hecho *la derivada* de la función $F(u, v) = (x, y)$ está dada justamente por la matriz jacobiana 2×2 cuyo determinante es el que aparece en la fórmula citada. Algo menos natural es que ahora, en la fórmula del cambio de variables en la integral doble, aparezca no sólo el determinante jacobiano, sino su *valor absoluto*. Esta situación se puede explicar en términos de la convención establecida para integrales simples de que $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$, de modo que se puede ver que en realidad, también en el caso del cambio de variable en una integral simple, se considera el valor absoluto del "jacobiano de la transformación $x = g(t)$" (que en este caso es $g'(t)$). No entraremos en detalles a este respecto.

Por otra parte, al igual que en el caso de la fórmula de cambio de variable en una integral simple, donde se pide que la función que define el cambio de variable $x = g(t)$ sea de clase $\mathscr{C}^1$, en el teorema 6.4.1 se pide también que la función $F(u, v) = (x, y) = (x(u, v), y(u, v))$ que define los cambios de variables $x = x(u, v)$, $y = y(u, v)$ sea de clase $\mathscr{C}^1$. Esto significa que las funciones componentes de F, $x = x(u, v)$, $y = y(u, v)$ son de clase $\mathscr{C}^1$, lo cual a su vez significa que las derivadas parciales $\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}, \frac{\partial y}{\partial u}, \frac{\partial y}{\partial v}$ de estas funciones componentes de F son continuas. Observe que con esta hipótesis aseguramos que el nuevo integrando que se obtiene al cambiar las variables, $f(F(u, v))\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$, es una función continua y por lo tanto integrable. Otro punto importante a observar es que en el teorema 6.4.1 se pide que la función de transformación de coordenadas F sea inyectiva. Este hecho nos permite que de las funciones coordenadas de F, $x = x(u, v)$, $y = y(u, v)$ podamos resolver para u, v en términos de x, y, obteniendo así las funciones $u = u(x, y)$, $v = v(x, y)$; es decir, obtener la función inversa $F^{-1}\colon R \subset \mathbb{R}^2 \to \mathbb{R}^2$, $F^{-1}(x, y) = (u, v) = (u(x, y), v(x, y))$.

Antes de continuar con el aspecto geométrico de la fórmula del teorema 6.4.1, veamos un primer ejemplo que muestra cómo trabaja esta fórmula.

Ejemplo 1. Retomemos el ejemplo 15 de la sección anterior. Se quiere calcular la integral de la función $f(x, y) = x + y + 1$ sobre la región R limitada por las rectas $y - x = 1$, $y - x = -1$, $y + x = 1$, $y + x = 2$. Esta región es un paralelogramo (de hecho un rectángulo) de vértices en $A = (0, 1)$, $B = (1, 0)$, $C = (3/2, 1/2)$, $D = (1/2, 3/2)$. Las ecuaciones de las rectas que limitan la región R sugieren la introducción de las nuevas variables u, v de modo que

$$u = y - x, \qquad v = y + x$$

En tal caso observe que el parelelogramo original se convierte en el rectángulo $-1 \leq u \leq 1$, $1 \leq v \leq 2$. Es decir, en las nuevas variables u y v la región de integración R' es el rectángulo mencionado. La función $F(u, v) = (x, y)$ de transformación de coordenadas es (despejando x, y de las ecuaciones $u = y - x$, $v = y + x$)

$$F(u, v) = (x, y) = \left(\frac{v - u}{2}, \frac{v + u}{2}\right)$$

El esquema se aprecaia en la figura 2.

En este caso, el jacobiano de la transformación es

$$\frac{\partial(x, y)}{\partial(u, v)} = \det\begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix} = \det\begin{bmatrix} -\dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{1}{2} \end{bmatrix} = -\frac{1}{2}$$

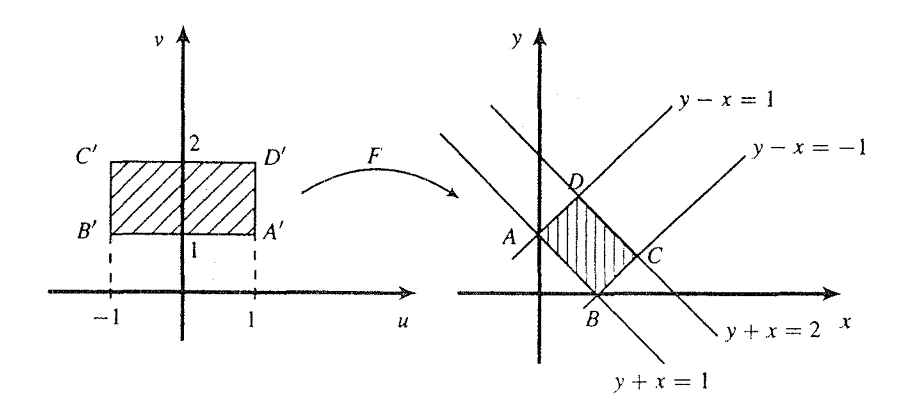


Figura 2. Las regiones R y R' del ejemplo 1

Haciendo pues nuestro cambio de variables nos queda que

$$\begin{aligned}\iint\limits_R f(x, y)\, dx\, dy &= \iint\limits_{R'} f(x(u, v), y(u, v))\left|\frac{\partial(x, y)}{\partial(u, v)}\right| du dv \\ &= \iint\limits_{R'}\left(\frac{v-u}{2} + \frac{v+u}{2} + 1\right)\left|-\frac{1}{2}\right| du dv \\ &= \frac{1}{2}\int_{-1}^{1}\int_{1}^{2}(v+1)\, dv du = \frac{1}{2}\int_{-1}^{1}\left[\frac{1}{2}(v+1)^2\right]_1^2 du \\ &= \frac{5}{4}\int_{-1}^{1} du = \frac{5}{4}(2) = \frac{5}{2}\end{aligned}$$

Este es el mismo resultado al que se llegó en el ejemplo 15 de la sección anterior. Lo que debe quedar claro es que el proceso para llegar a él fue notablemente más sencillo y económico. ■

Aunque no presentaremos aquí la prueba del teorema 6.4.1, sí veremos ahora un argumento geométrico que nos deja a nivel de "aceptable" el resultado en el caso particular en que la función $f(x, y)$ es constante e igual a 1. Es decir, presentaremos un argumento geométrico que hace plausible la fórmula

$$\iint\limits_R dx\, dy = \iint\limits_{R'}\left|\frac{\partial(\phi, \psi)}{\partial(u, v)}\right| du dv$$

Usaremos algunos de los conceptos de curvas estudiados en el capítulo anterior. Consideremos la región R' en el plano uv dada por

$$R' = \{(u, v)|u_0 \le u \le u_0 + \Delta u,\ v_0 \le v \le v_0 + \Delta v\}$$

Esta región es entonces un rectángulo con vértices en los puntos $A' = (u_0, v_0)$, $B' = (u_0 + \Delta u, v_0)$, $C' = (u_0 + \Delta u, v_0 + \Delta v)$, $D' = (u_0, v_0 + \Delta v)$. Al aplicar la transformación $F\colon R' \to \mathbb{R}^2$, la región R'

se transforma en la región $R = F(R')$. Estaremos interesados en comparar las áreas de las regiones R y R'; más aún, puesto que el área de R' es $\Delta u \Delta v$, estaremos interesados en calcular el área de la región R y compararla con la de R'. De hecho, si pensamos en un factor K (que depende del punto (u, v) en cuestión) de "modificación de áreas", tal que (área de R) $= K$(área de R'), veremos que este factor es ni más ni menos que el jacobiano de la transformación.

Sea entonces $F: R' \to \mathbb{R}^2$, $F(u, v) = (x, y) = (x(u, v), y(u, v))$ la función de transformación de coordenadas, que manda la región R' del plano uv en la región R del plano xy. Si Δu y Δv son pequeños, el rectángulo R' se transformará por medio de F, en una región R "parecida a un rectángulo", (un "paralelogramo curvilíneo") con vértices en $A = F(A')$, $B = F(B')$, $C = F(C')$, $D = F(D')$. En un esquema se ve como

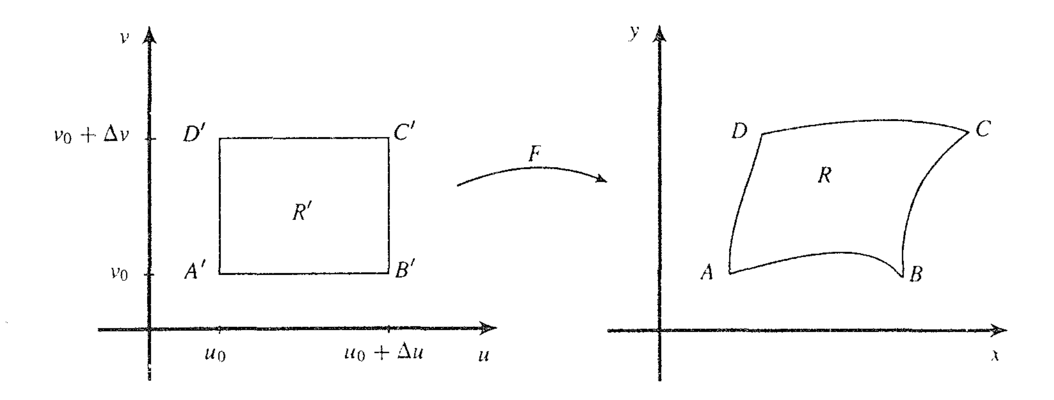


Figura 3. La transformación del rectángulo R' en el "rectángulo" R por medio de la función $F(u, v) = (x, y)$

Consideremos la *curva* $\alpha: [u_0, u_0 + \Delta u] \to \mathbb{R}^2$, dada por

$$\alpha(t) = F(t, v_0) = (x(t, v_0), y(t, v_0))$$

Nótese que la traza de esta curva es el "lado" AB del paralelogramo curvilíneo R. Considere también la curva $\beta: [v_0, v_0 + \Delta v] \to \mathbb{R}^2$, dada por

$$\beta(t) = F(u_0, t) = (x(u_0, t), y(u_0, t))$$

La traza de esta curva es el "lado" AD del paralelogramo curvilíneo R. Los vectores velocidad de estas curvas son

$$\alpha'(t) = \left(\frac{\partial x}{\partial u}(t, v_0), \frac{\partial y}{\partial u}(t, v_0)\right), \quad \beta'(t) = \left(\frac{\partial x}{\partial v}(u_0, t), \frac{\partial y}{\partial v}(u_0, t)\right)$$

Las longitudes de los lados AB y AD de la región R se pueden calcular entonces como

$$L_{AB} = \int_{u_0}^{u_0+\Delta u} \|\alpha'(t)\|\, dt, \quad L_{AD} = \int_{v_0}^{v_0+\Delta v} \|\beta'(t)\|\, dt$$

Si Δu y Δv son pequeños, estas integrales son casi iguales a (usando el teorema del valor medio) $\|\alpha'(u_0)\|\Delta u$ y $\|\beta'(v_0)\|\Delta v$, en forma respectiva, las cuales a su vez son las normas de los vectores

$$U = \alpha'(u_0)\Delta u \quad \text{y} \quad V = \beta'(v_0)\Delta v$$

o bien

$$U = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}\right)\Delta u \quad \text{y} \quad V = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}\right)\Delta v$$

donde todas las derivadas están evaluadas en (u_0, v_0). En esquema tenemos

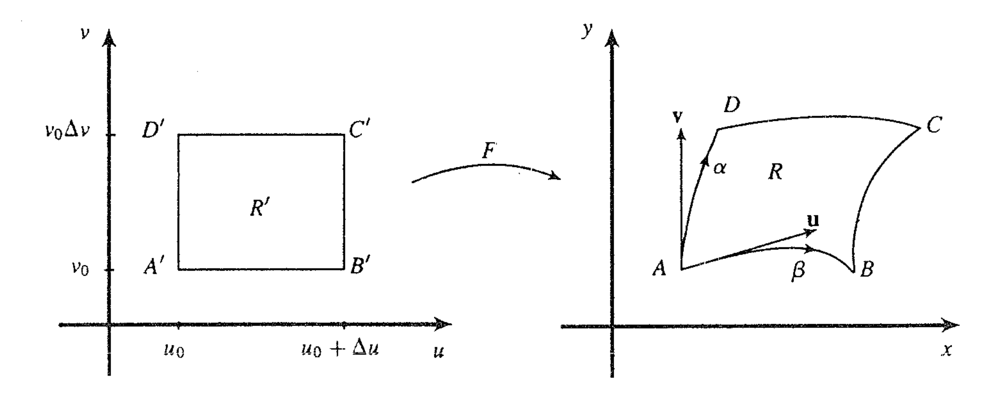


Figura 4. El rectángulo R' se transforma en el paralelogramo curvilíneo R cuyos lados AB y AD son aproximadamente las normas de los vectores U y V respectivamente.

El área del paralelogramo curvilíneo R es entonces aproximadamente igual al área del paralelogramo generado por los vectores U y V. Esta aproximación será tanto mejor en cuanto Δu y Δv sean pequeños. Sabemos que el área del paralelogramo generado por dos vectores es igual a la norma del producto cruz de estos vectores (ver discusión previa al ejemplo 3 de la sección 5 del capítulo 1), que como vectores de $\mathbb{R}^3$, tienen su última coordenada igual a cero. Entonces el área de la región R es aproximadamente igual a la norma del vector

$$\begin{aligned} U \times V &= (\alpha'(u_0)\Delta u) \times (\beta'(v_0)\Delta v) = \Delta u \Delta v \alpha'(u_0) \times \beta'(v_0) \\ &= \Delta u \Delta v \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & 0 \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & 0 \end{bmatrix} \\ &= \Delta u \Delta v \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \mathbf{k} = \Delta u \Delta v \frac{\partial(x, y)}{\partial(u, v)}\mathbf{k} \end{aligned}$$

de modo que el área de la región R es, aproximadamente

$$\|\alpha'(u_0) \times \beta'(v_0)\| = \left|\frac{\partial(x, y)}{\partial(u, v)}\right| \Delta u \Delta v$$

En resumen, se tiene que

$$\text{área de } R = \left|\frac{\partial(x, y)}{\partial(u, v)}\right| \text{ (área de } R')$$

Antes de terminar este tema, veamos un ejemplo.

Ejemplo 2. Retomemos la situación presentada en el ejemplo 1 con las regiones R' y R de integración. La región R del plano xy sobre la que se quiere calcular la integral es el paralelogramo de vértices en $A = (0, 1)$, $B = (1, 0)$, $C = (3/2, 1/2)$, $D = (1/2, 3/2)$. Esta región correspondía, bajo la transformación $F: R' \to \mathbb{R}^2$, $F(u, v) = (x, y) = \left(\frac{v-u}{2}, \frac{v+u}{2}\right)$, al rectángulo en el plano uv, $R = \{(u, v) | -1 \leq u \leq 1, 1 \leq v \leq 2\}$. Observe que el área de la región R es 1 (se trata de un rectángulo de lados $|\overline{AB}| = \sqrt{2}$, $|\overline{BC}| = 1/\sqrt{2}$, y el área de la región R' es 2 (un rectángulo en el plano uv de lados 2 y 1), de modo que

$$\frac{\text{área de } R}{\text{área de } R'} = \frac{1}{2}$$

Este valor del cociente de las dos áreas es justamente el valor que en el ejemplo 1 calculamos del (valor absoluto del) jacobiano de la transformación. Es decir, se tiene

$$\frac{\text{área de } R}{\text{área de } R'} = \frac{1}{2} = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|$$

como habíamos establecido previamente a este ejemplo. (**NOTA:** en este ejemplo particular las regiones R' y R son rectángulos y el jacobiano es *constante*. Esto, en general, no ocurre: el jacobiano es un factor *local* de modificación de áreas). ■

Consideremos ahora, para finalizar esta exposición, un rectángulo cualquiera $Q = [a, b] \times [c, d]$ (en el plano uv) y tomemos una partición P de él en subrectángulos "pequeños" R'_{ij} como el de la región R' estudiada anteriormente. Podemos considerar que si estos subrectángulos son realmente pequeños, la función $g(u, v) = \left|\frac{\partial(x,y)}{\partial(u,v)}\right|$ se mantiene constante en cada uno de ellos, de tal manera que podemos pensar en ella como una función escalonada en Q. Su integral doble sobre Q se calcularía como

$$\iint_Q g(u, v)\, du dv = \sum_i \sum_j (\text{valor de } g(u, v) \text{ en el rectángulo } R'_{ij})(\text{área de } R'_{ij})$$

$$= \sum_i \sum_j \left|\frac{\partial(x, y)}{\partial(u, v)}\right| (\text{área de } R'_{ij}) = \sum_i \sum_j \text{área de } R_{ij} = \text{área de } F(Q)$$

donde, como antes, estamos llamando R_{ij} al paralelogramo curvilíneo (en el plano xy) de imagen bajo la transformación F del rectángulo R'_{ij}. Denotando por R la imagen bajo F de Q (es decir, $R = F(Q)$, el cual ya puede ser "muy diferente" de un paralelogramo curvilíneo, pues ahora Q es

un rectángulo "grande"), y observando que el área de R se puede calcular como la integral doble de la función $f(x, y) = 1$ sobre la región R (este es un hecho que se expondrá con calma en la siguiente sección; por el momento trate de ver esta afirmación desde del punto de vista geométrico: el volumen bajo la superficie $z = 1$ sobre la región R es el área de R por la altura que es constante igual a 1, es decir, es numéricamente igual al área de R), se tiene que

$$\iint_Q g(u, v)\, du dv = \iint_Q \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du dv = \text{área de } R = \iint_{F(Q)} dx\, dy$$

o sea

$$\iint_{F(Q)} dx\, dy = \iint_Q \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du dv$$

que es la fórmula del teorema 6.4.1 en el caso en que $f(x, y) = 1$.

Veamos algunos ejemplos.

Ejemplo 3. Queremos calcular la integral de la función $f(x, y) = x^2y^2$ sobre la región R limitada por las hipérbolas equiláteras $xy = 1$, $xy = 2$ y las rectas $y = \frac{x}{2}$, $y = 3x$ (la región situada en el primer cuadrante). Tanto la función a integrar como la región de integración sugieren el cambio de variables $xy = u$, $\frac{y}{x} = v$. Según este cambio, la región de integración corresponde al rectángulo (del plano uv) $1 \leq u \leq 2$, $0.5 \leq v \leq 3$. Obsérvese que de hecho no decidimos el cambio de variables con las expresiones explícitas de las funciones $x = x(u, v)$, $y = y(u, v)$, coordenadas de la que hemos venido llamando función de transformación de coordenadas $F: R \to \mathbb{R}^2$, sino *con la inversa de esta función*. Nótese también que el jacobiano que necesitamos calcular es el de la función F, de modo que: 1) hacemos explícitas las funciones $x = x(u, v)$, $y = y(u, v)$, ó 2) usamos la propiedad estudiada en el capítulo 3 de que la matriz jacobiana de F es la inversa de la correspondiente matriz de F^{-1} (esta última se calcula fácilmente de las fórmulas establecidas para $u = u(x, y) = xy$, $v = v(x, y) = y/x$), y por lo tanto su determinante (jacobiano) es el inverso (algebraico) del jacobiano de F^{-1}. Hagámoslo por esta vez de las dos maneras.

Si $u = xy$, $v = y/x$, nos queda, resolviendo para x, y en términos de u, v las siguientes fórmulas (tomando en cuenta que solamente trabajaremos en los primeros cuadrantes de los planos x, y y u, v)

$$x = x(u, v) = (u/v)^{1/2}, \qquad y = y(u, v) = (uv)^{1/2}$$

de modo que el jacobiano de la transformación $F(u, v) = (x, y)$ es

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \det \begin{bmatrix} 0.5(uv)^{-1/2} & -0.5u^{1/2}v^{-3/2} \\ 0.5(v/u)^{1/2} & 0.5(u/v)^{1/2} \end{bmatrix} = \frac{1}{2v}$$

De otro modo, si calculamos el jacobiano $\left|\frac{\partial(u,v)}{\partial(x,y)}\right|$ con las funciones establecidas desde el principio $u = xy$, $v = y/x$, nos queda

$$\frac{\partial(u, v)}{\partial(x, y)} = \det \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \det \begin{bmatrix} y & x \\ -y/x^2 & 1/x \end{bmatrix} = \frac{2y}{x} = 2v$$

de modo que el jacobiano que buscamos es

$$\frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{\dfrac{\partial(u, v)}{\partial(x, y)}} = \frac{1}{2v}$$

Se tiene pues en esquema que

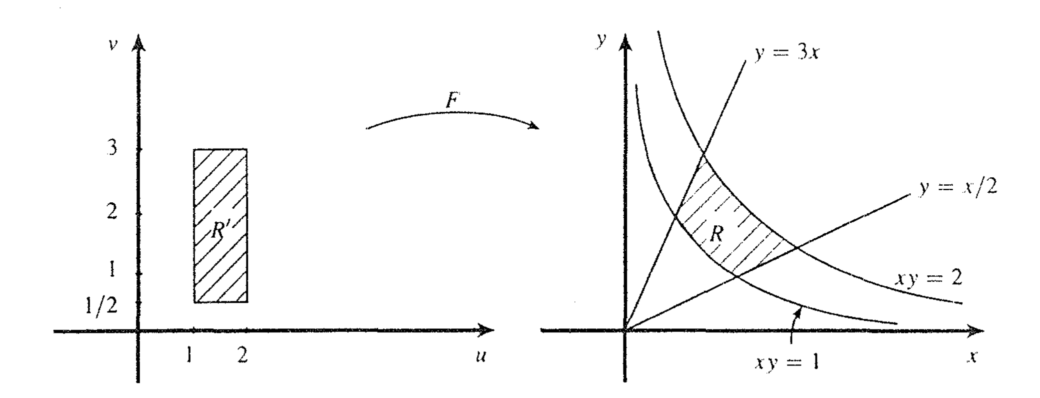


Figura 5. La transformación de coordenadas para el ejemplo 3.

La función a integrar es, en términos de las nuevas variables, $f(F(u, v)) = u^2$. La fórmula cambio de variables se ve en este caso como

$$\iint_R x^2y^2\, dx\, dy = \iint_{R'} u^2 \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du dv = \int_1^2 \int_{0.5}^3 u^2\left(\frac{1}{2v}\right) dv du$$

$$= \int_1^2 u^2\, du \int_{0.5}^3 \frac{1}{2v}\, dv = \int_1^2 u^2\, du\left(\frac{1}{2}\ln 6\right) = \frac{7}{6}\ln 6$$

(Sugerimos al lector que haga el cálculo de la integral de este ejemplo sin recurrir al cambio de variables; ¡esperamos que note la diferencia!). ■

Ejemplo 4. Queremos calcular la integral de la función $f(x, y) = x^{-3}$ sobre la región limitada por las parábolas $y = x^2$, $y = 2x^2$, $y^2 = x$, $y^2 = 2x$. Dejamos al lector que se convenza de que la evaluación directa de esta integral nos conduciría a realizar los siguientes cálculos

$$\int_{4^{-1/3}}^{2^{-1/3}} \int_{x^{1/2}}^{2x^2} x^{-3} dydx + \int_{2^{-1/3}}^{1} \int_{x^{1/2}}^{(2x)^{1/2}} x^{-3}\, dydx + \int_1^{2^{1/3}} \int_{x^2}^{(2x)^{1/2}} x^{-3}\, dydx$$

Sin embargo, haciendo el cambio de variables $u = y/x^2$, $v = y^2/x$, la región de integración R corresponde al cuadrado $R' = \{(u, v)|1 \leq u, v \leq 2\}$, la función a integrar es $f(F(u, v)) = u^2/v$, y

el jacobiano, calculado como el recíproco del jacobiano de u, v respecto de x, y es

$$\frac{\partial(x, y)}{\partial(u, v)} = \left(\frac{\partial(u, v)}{\partial(x, y)}\right)^{-1} = \left(\det\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}\right)^{-1}$$
$$= \left(\det\begin{bmatrix} -2y/x^3 & 1/x^2 \\ -y^2/x^2 & 2y/x \end{bmatrix}\right)^{-1} = \left(\frac{-3y^2}{x^4}\right)^{-1}$$
$$= \left(-3u^2\right)^{-1} = -\frac{1}{3u^2}$$

de modo que el cálculo de la integral se reduce a

$$\iint_R x^{-3}\, dx\, dy = \iint_{R'} u^2/v \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du dv = \int_1^2 \int_1^2 (u^2/v)(1/3u^2)\, dv du$$
$$= \int_1^2 \int_1^2 \frac{1}{3v}\, dv du = \frac{1}{3}\ln 2$$

■

En esta última parte de la presente sección, vamos a considerar un caso particular de cambio de variables en integrales dobles que es muy importante: el cambio a coordenadas polares. Recuerde que en el sistema coordenado polar se localizan los puntos en el plano por medio de dos parámetros, a saber: la distancia $r \geq 0$ del punto al origen (llamado polo) y el ángulo θ que forma el radio vector del punto (el vector—como segmento dirigido—correspondiente al punto) con la parte positiva del eje x. De esta manera decimos que el punto P tiene por coordenadas polares la pareja (r, θ). Llamaremos entonces r, θ a las nuevas coordenadas, en lugar de u, v. La equivalencia entre estas coordenadas y las coordenadas cartesianas (x, y) del punto P es, como sabemos

$$x = r\cos\theta, \quad y = r \operatorname{sen}\theta$$

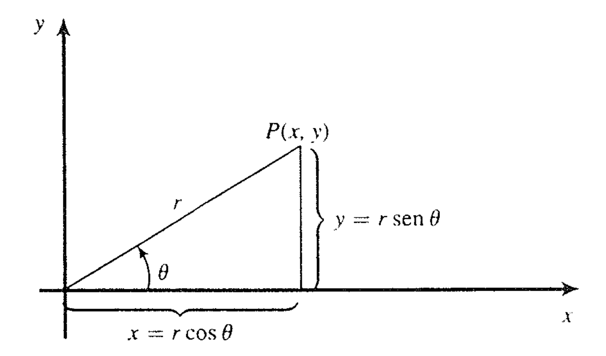


Figura 6. Correpondencia entre las cooerdenadas cartesianas (x, y) de un punto P en el plano y sus coordenadas polares (r, θ)

Las expresiones anteriores son, de hecho, las funciones coordenadas de la que hemos venido llamando función de transformación de coordenadas $F: R' \to \mathbb{R}^2$, $F(r, \theta) = (x, y) = (x(r, \theta), y(r, \theta)) = (r\cos\theta, r\,\mathrm{sen}\,\theta)$. Es importante hacer notar que para que esta función trabaje bien (en el aspecto de inyectividad), se debe considerar a la coordenada θ variando solamente entre 0 y 2π. Así, el plano $r\theta$ se describe con los rangos de variación $r \geq 0$, $0 \leq \theta < 2\pi$. Un hecho de fundamental importancia en este cambio de variables es notar que las curvas del tipo $r =$ cte. (en el plano $r\theta$) corresponden a círculos del tipo $x^2 + y^2 = (\text{cte.})^2$. En efecto, si $r = r_0$, se tiene que $x^2 + y^2 = (r_0\cos\theta)^2 + (r_0\,\mathrm{sen}\,\theta)^2 = r_0^2$ que es la ecuación de un círculo con centro en el origen y radio r_0. También se puede ver que las curvas del tipo $\theta = \theta_0$ en el plano $r\theta$ corresponden a rectas radiales en el plano xy (de hecho, a las rectas $y = (\tan\theta_0)x$, si $\theta_0 \neq \pi/2, 3\pi/2$, en cuyo caso es la recta $x = 0$). Así pues, los cuadriláteros curvilíneos limitados por sectores circulares y rectas radiales como

$$R = \{(x, y) | r_0^2 \leq x^2 + y^2 \leq r_1^2, (\tan\theta_0)x \leq y \leq (\tan\theta_1)x\}$$

corresponden en el plano $r\theta$ al rectángulo

$$R' = \{(r, \theta) | r_0 \leq r \leq r_1, \theta_0 \leq \theta \leq \theta_1\}$$

como se muestra esquemáticamente en la figura siguiente

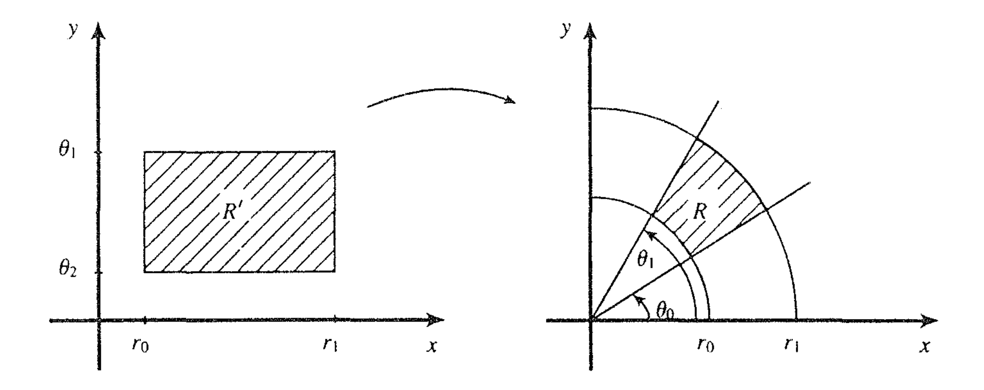


Figura 7. La transformación $F(r, \theta) = (x, y) = (r\cos\theta, r\,\mathrm{sen}\,\theta)$.

En este caso, el jacobiano de la transformación que aparece en la fórmula de cambio de variables es

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det\begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial\theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial\theta} \end{bmatrix} = \det\begin{bmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{bmatrix} = r$$

(nótese que el jacobiano de la transformación se anula en $r = 0$. Esto, sin embargo, no anula la validez del teorema 6.4.1 pues el conjunto $\{(r, \theta) | r = 0\}$ es de medida cero). Entonces la fórmula

del teorema 6.4.1 se ve (en este caso particular) como

$$\iint_R f(x,y)\,dx\,dy = \iint_{R'} f(x(r,\theta), y(r,\theta))\left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| dr d\theta$$

$$= \iint_{R'} f(r\cos\theta, r\operatorname{sen}\theta) r dr\, d\theta$$

Aunque no hay un criterio universal que nos diga cuándo debemos hacer el cambio a coordenadas polares al calcular una integral doble de una función, sí podemos decir que, en general, cuando en la región de integración se presenten anillos circulares (o trozos de ellos), y/o cuando en la función $f(x, y)$ a integrar aparezca de alguna forma las expresiones $(x^2 + y^2)$ y/o $\frac{y}{x}$ (que en coordenadas polares corresponden a r^2 y a $(\tan\theta)$, respectivamente), puede resultar conveniente intentar el cálculo de la integral haciendo previamente el cambio a coordenadas polares. A continuación se estudian algunos ejemplos típicos.

Ejemplo 5. Retomemos el ejemplo 13 de la sección anterior. Se quiere calcular la integral $\iint_R f(x, y)\,dx\,dy$ de la función

$$f(x,y) = \sqrt{R^2 - x^2 - y^2}$$

en que la región R es $x^2 + y^2 \leq R^2$. Este es un típico caso en el que, bajo cualquier punto de vista conviene hacer el cambio a coordenadas polares para evaluar la integral. Obsérvese que la región de integración R se describe, en coordenadas polares por $0 \leq r \leq R$, y $0 \leq \theta \leq 2\pi$. De modo que la región de integración R' se ve (en el plano $r\theta$) como el rectángulo $0 \leq r \leq R$, $0 \leq \theta \leq 2\pi$. Así pues

$$\iint_R f(x,y)\,dx\,dy = \int_{-R}^{R}\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \sqrt{R^2 - x^2 - y^2}\,dy\,dx = \int_0^R \int_0^{2\pi} \sqrt{R^2 - r^2} r\,dr\,d\theta$$

$$= \int_0^R \sqrt{R^2 - r^2} r\,dr \int_0^{2\pi} d\theta = \frac{-1}{3}(R^2 - r^2)^{3/2}\bigg|_0^R (2\pi) = \frac{2\pi}{3} R^3$$

resultado que habíamos obtenido ya en la sección anterior, calculando directamente la integral en el sistema xy. ■

Ejemplo 6. Retomemos el ejemplo 14 de la sección anterior. Se quiere calcular la integral $\iint_R (x+y)\,dx\,dy$ donde R es la cuarta parte, situada en el primer cuadrante, del anillo circular limitado por los círculos $x^2 + y^2 = 1$, $x^2 + y^2 = 4$. En coordenadas polares esta región de integración se describe con r que varía de 1 a 2 (las ecuaciones de los dos círculos son $r = 1$, $r = 2$, y se está considerando la región comprendida entre ellos) y θ varía de 0 a $\pi/2$ (así cubrimos el primer cuadrante). Se tiene entonces el rectángulo $R' = \{(r,\theta) | 1 \leq r \leq 2, 0 \leq \theta \leq \pi/2\}$. Haciendo las sustituciones correspondientes nos queda

$$\iint_R (x+y)\,dx\,dy = \int_1^2 \int_0^{\pi/2} (r\cos\theta + r\operatorname{sen}\theta) r\,dr d\theta$$

$$= \int_1^2 r^2\, dr \int_0^{\pi/2} (\cos\theta + \operatorname{sen}\theta)\, d\theta = \int_1^2 r^2\, dr(2) = (2)\left(\frac{1}{3}\right) r^3 \Big]_1^2$$
$$= \frac{14}{3}$$

como obtuvimos en la sección anterior (con mucho más trabajo). ∎

Ejemplo 7. Sea calcular la integral de la función $f(x, y) = \ln(1 + x^2 + y^2)$ sobre la región R limitada por el círculo unitario $x^2 + y^2 = 1$. Al hacer el cambio a coordenadas polares la región de integración corresponde a $R' = \{(r, \theta)|0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}$ y la función a integrar es $f(r\cos\theta, r\operatorname{sen}\theta) = \ln(1 + r^2)$. Se tiene entonces que

$$\iint_R \ln(1 + x^2 + y^2)\, dx\, dy = \int_0^1 \int_0^{2\pi} \ln(1 + r^2) r\, dr d\theta$$
$$= 2\pi \int_0^1 \ln(1 + r^2) r\, dr = \pi(1 + r^2)\big(\ln(1 + r^2) - 1\big)\Big]_0^1$$
$$= \pi(2\ln 2 - 1)$$ ∎

Ejemplo 8. (El área bajo la curva de la campana de Gauss). Consideremos la función $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = e^{-x^2}$. Esta es una función de amplio uso en la teoría de las probabilidades, cuya gráfica es conocida como la "campana de Gauss". En muchos problemas importantes (de probabilidad) resulta indispensable calcular el área bajo esta curva entre $x = 0$ y algún otro valor $x = x_0$. Es decir, se trata de calcular la integral

$$\int_0^{x_0} e^{-x^2}\, dx$$

que corresponde al cálculo del área marcada en la siguiente figura

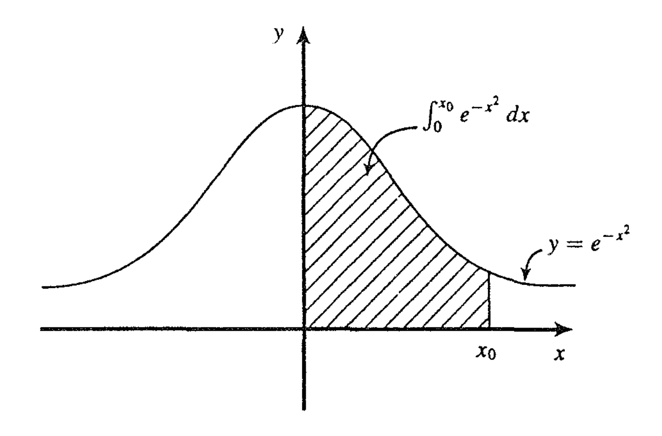


Figura 8. La campana de Gauss.

Según el teorema fundamental del cálculo, para poder hacer éste, debemos conocer una antiderivada de e^{-x^2}. Y aquí es donde se nos acaba la alegría por resolver el problema: se sabe que NO EXISTE

función alguna escrita en términos de funciones elementales cuya derivada sea e^{-x^2} (en palabras más llanas: la integral de e^{-x^2} *no se puede resolver*). Por eso el problema tiene que ser resuelto numéricamente y encontramos en los apéndices de muchos libros de probabilidad tablas con los valores de esta integral para diferentes valores de x_0. Un hecho sorprendente es que aunque no se pueda calcular la integral de la función e^{-x^2} entre 0 y x_0, *con* x_0 *finito*, sí se puede calcular con exactitud la integral impropia de esta función entre $x = 0$ e infinito. Es decir, se puede calcular con toda exactitud la integral impropia

$$I = \int_0^{\infty} e^{-x^2}\, dx$$

que corresponde al cálculo de la mitad del área total bajo la campana de Gauss, es decir, a la mitad del área limitada por el eje de las x y la campana de Gauss. Para hacer este cálculo se recurre a un ingenioso truco que consiste en calcular, pasando a las coordenadas polares, la integral doble definida como el producto de la integral I por ella misma. Más en concreto, se tiene

$$I^2 = \left(\int_0^{\infty} e^{-x^2}\, dx\right)\left(\int_0^{\infty} e^{-y^2}\, dy\right) = \int_0^{\infty}\int_0^{\infty} e^{-(x^2+y^2)}\, dx\, dy$$

Obsérvese que esta última integral se está haciendo *sobre todo el primer cuadrante del plano* xy, el cual, descrito en coordenadas polares, corresponde a una variación en r de 0 a inifinito y en θ de 0 a $\pi/2$. Así pues, escribiendo esta integral en coordenadas polares se tiene que

$$I^2 = \int_0^{\infty}\int_0^{\pi/2} e^{-r^2} r\, dr d\theta = -\frac{\pi}{4}\int_0^{\infty} e^{-r^2}(-2r)\, dr = -\frac{\pi}{4}e^{-r^2}\Big]_0^{\infty} = \frac{\pi}{4}$$

de modo que la integral I vale

$$I = \int_0^{\infty} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}$$

■

Un segundo cambio de variables a estudiar es el llamado "cambio a coordenadas polares generalizadas", en el cual las nuevas coordenadas r, θ están relacionadas con las coordenadas x, y como

$$x = ar\cos\theta, \quad y = br\operatorname{sen}\theta$$

Por ejemplo, en este sistema coordenado la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ corresponde a la recta $r = 1$. El jacobiano de la transformación, en este caso, es

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det\begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta}\end{bmatrix} = \det\begin{bmatrix} a\cos\theta & -ar\operatorname{sen}\theta \\ b\operatorname{sen}\theta & br\cos\theta\end{bmatrix} = abr$$

de modo que la fórmula de cambio de variables se presenta como

$$\iint_R f(x, y)\, dx\, dy = ab\iint_{R'} f(ar\cos\theta, br\operatorname{sen}\theta) r\, dr d\theta$$

donde R' es la región del plano $r\theta$ que bajo la transformación de coordenadas corresponde a la región R del plano xy.

Ejemplo 9. Queremos calcular la integral de la función

$$f(x, y) = \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}$$

sobre la región R limitada por la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Haciendo el cambio a coordenadas polares generalizadas $x = ar\cos\theta$, $y = br\,\mathrm{sen}\,\theta$, la región de integración se describe como $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$, de modo que

$$\begin{aligned}\iint_R \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\,dx\,dy &= \int_0^1 \int_0^{2\pi} \sqrt{1 - r^2}abr\,dr d\theta \\ &= 2\pi ab \int_0^1 \sqrt{1 - r^2}r\,dr = 2\pi ab\left(-\frac{1}{3}(1 - r^2)^{3/2}\right)\Big|_0^1 \\ &= \frac{2}{3}\pi ab\end{aligned}$$

(Si el lector resolvió el ejercicio 31 de la sección anterior, debe notar la gran ventaja que proporciona hacer el cambio de variables en la integral doble). ■

Ejemplo 10. Se quiere calcular la integral de la función $f(x, y) = 1$ sobre el anillo elíptico limitado por las elipses $x^2 + 4y^2 = 4$, $x^2 + 4y^2 = 16$. Haciendo el cambio a coordenadas polares generalizadas, según las ecuaciones $x = 2r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$, las elipses corresponden a las rectas $r = 1$ y $r = 2$. Entonces, la región de integración se describe en este nuevo sistema como el rectángulo $R' = \{(r, \theta)|1 \leq r \leq 2, 0 \leq \theta \leq 2\pi\}$, y la integral procurada es

$$\int_1^2 \int_0^{2\pi} (2)(1)(r)\,dr d\theta = 2(2\pi)\int_1^2 r\,dr = 6\pi$$ ■

Se puede hacer incluso una generalización del cambio de variables a coordenadas polares generalizadas, escribiendo

$$x = ar\cos^\alpha\theta, \quad y = br\,\mathrm{sen}^\alpha\theta$$

En este caso el jacobiano de la transformación es

$$\begin{aligned}\frac{\partial(x, y)}{\partial(r, \theta)} &= \det\begin{bmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{bmatrix} \\ &= \det\begin{bmatrix} a\cos^\alpha\theta & -a\alpha r\cos^{\alpha-1}\theta\,\mathrm{sen}\,\theta \\ b\,\mathrm{sen}^\alpha\theta & b\alpha r\,\mathrm{sen}^{\alpha-1}\theta\cos\theta\end{bmatrix} = ab\alpha r\,\mathrm{sen}^{\alpha-1}\theta\cos^{\alpha-1}\theta\end{aligned}$$

El siguiente ejemplo muestra cómo este cambio de variables puede ser útil al calcular una integral doble

Ejemplo 11. Se quiere integrar la función constante $f(x, y) = 1$ sobre la región limitada por la curva

$$\left(\frac{x}{a} + \frac{y}{b}\right)^4 = \frac{x^2}{h^2} + \frac{y^2}{k^2}$$

en el primer cuadrante (a, b, h, k son números positivos dados). En este caso se presenta una situación interesante: el cambio de variables no sólo es conveniente para hacer la integral, sino que es *necesario*. De otro modo, estaríamos obligados a hacer explícita la función $y = y(x)$ que define el "techo" de nuestra región de integración, lo que nos conduciría a resolver una ecuación de cuarto grado para la variable y, ecuación que deberíamos integrar posteriormente respecto de x. Esto, por supuesto, plantea un panorama poco agradable. Sin embargo, si introducimos las variables r y θ de modo que

$$x = ar\cos^2\theta, \qquad y = br\operatorname{sen}^2\theta$$

la ecuación que define nuestra región de integración se transforma en

$$\left(\frac{ar\cos^2\theta}{a} + \frac{br\operatorname{sen}^2\theta}{b}\right)^4 = \frac{a^2r^2\cos^4\theta}{h^2} + \frac{b^2r^2\operatorname{sen}^4\theta}{k^2}$$

o bien, simplificando

$$r = \left(\frac{a^2}{h^2}\cos^4\theta + \frac{b^2}{k^2}\operatorname{sen}^4\theta\right)^{1/2}$$

Esta es una función *explícita* del tipo $r = r(\theta)$ que nos define el límite de nuestra región de integración. Así, debemos integrar en la nueva variable r de 0 a la expresión obtenida, y en θ de 0 a $\pi/2$ (para cubrir el primer cuadrante). En este caso el jacobiano de la transformación es

$$\frac{\partial(x, y)}{\partial(r, \theta)} = 2abr\cos\theta\operatorname{sen}\theta$$

de modo que la integral procurada es

$$\begin{aligned}
\iint_R f(x, y)\,dx\,dy &= \iint_R dx\,dy = \iint_{R'} 2abr\cos\theta\operatorname{sen}\theta\,dr d\theta \\
&= 2ab\int_0^{\pi/2}\int_0^{(\frac{a^2}{h^2}\cos^4\theta+\frac{b^2}{k^2}\operatorname{sen}^4\theta)^{1/2}} r\cos\theta\operatorname{sen}\,dr d\theta \\
&= ab\int_0^{\pi/2}\operatorname{sen}\theta\cos\theta\left(\frac{a^2}{h^2}\cos^4\theta + \frac{b^2}{k^2}\operatorname{sen}^4\theta\right)d\theta \\
&= ab\left[-\frac{a^2}{6h^2}\cos^6\theta + \frac{b^2}{6k^2}\operatorname{sen}^6\theta\right]_0^{\pi/2} = \frac{ab}{6}\left(\frac{a^2}{h^2} + \frac{b^2}{k^2}\right)
\end{aligned}$$

■

Ejercicios (Capítulo 6, Sección 4)

1. Considere un paralelogramo R limitado por las dos parejas de rectas

$$\begin{matrix} ax + by = \alpha \\ ax + by = \beta \end{matrix} \quad \text{y} \quad \begin{matrix} cx + dy = \gamma \\ cx + dy = \delta \end{matrix}$$

(donde $ad - bc \neq 0$). Suponga que se quiere calcular la integral doble de la función $f(x, y)$ sobre la región R. Introduciendo las nuevas variables u, v

$$u = ax + by, \quad v = cx + dy$$

describa la región R' sobre la que se integraría en el plano uv. Escriba explícitamente la fórmula de cambio de variables del teorema 6.4.1 en este caso particular.

2. Considere la región R limitada por las hipérbolas $xy = \alpha$, $xy = \beta$, y por las rectas $y = \gamma x$, $y = \delta x$, donde $\alpha, \beta, \gamma, \delta > 0$ (la región en el primer cuadrante). Suponga que se quiere calcular la integral doble de la función $f(x, y)$ sobre la región R. Introduciendo las nuevas variables u, v

$$u = xy, \quad v = y/x$$

describa la región R' sobre la que se integraría en el plano uv. Escriba explícitamente la fórmula de cambio de variables del teorema 6.4.1 en este caso particular.

En cada uno de los ejercicios 3—12, se dan regiones R en el plano cartesiano. Describa cada una de estas regiones en el sistema de coordenadas polares.

3. $R = \{(x, y)|x^2 + y^2 \leq 4, y \geq 0\}$

4. $R = \{(x, y)|x^2 + y^2 \leq 4, y \leq 0\}$

5. $R = \{(x, y)|1 \leq x^2 + y^2 \leq 9, x \geq 0\}$

6. $R = \{(x, y)|(x - 1)^2 + y^2 \leq 1\}$

7. $R = \{(x, y)|x^2 + (y - 2)^2 \leq 4\}$

8. $R = \{(x, y)|x^2 + y^2 + 4x \leq 0\}$

9. $R = \{(x, y)|x^2 + y^2 + 6y \leq 0\}$

10. R es la región en el primer cuadrante limitada por los círculos $x^2 + y^2 = 2$, $x^2 + y^2 = 3$, y las rectas $y = x$, $y = 2x$.

11. R es la región en el tercer cuadrante limitada por los círculos $x^2 + y^2 = 1$, $x^2 + y^2 = 3$, y las rectas $y = x$, $y = 3x$.

12. R es la región en el segundo cuadrante limitada por los círculos $x^2 + y^2 = 2$, $x^2 + y^2 = 4$, y las rectas $y = -x$, $y = 0$.

En los ejercicios 13—25, calcule la integral doble indicada, efectuando un cambio de variables adecuado.

13. $\iint_R (x + y)\,dx\,dy$, donde R es la región limitada por el cuadrado $|x| + |y| = 1$.

14. $\iint_R x\,dx\,dy$, donde R es la región limitada por el paralelogramo cuyos vértices son $A = (0, 0)$, $B = (6, 2)$, $C = (10, 5)$, $D = (4, 3)$.

15. $\iint_R y\,dx\,dy$, donde R es la región limitada por el paralelogramo cuyos lados son $y = x - 2$, $y = x + 1$, $y = 2 - x$, $y = -x$.

16. $\iint_R (x^2 + y^2)\,dx\,dy$, donde $R = \{(x, y)|x^2 + y^2 \leq 4\}$.

17. $\iint_R \exp(x^2 + y^2)\,dx\,dy$, donde $R = \{(x, y)|x^2 + y^2 \leq 9\}$

18. $\iint_R \frac{1}{4 - x^2 - y^2}\,dx\,dy$, donde $R = \{(x, y)|x^2 + y^2 \leq 1\}$

19. $\iint_R (x + y)^2\,dx\,dy$, donde R es la región limitada por el círculo de centro en $(2, 0)$, tangente al eje y.

20. $\iint_R \arctan\frac{y}{x}\,dx\,dy$, donde R es la región en el primer cuadrante, limitada por los círculos $x^2 + y^2 = 1$, $x^2 + y^2 = 4$, y las rectas $y = x$, $y = \sqrt{3}x$.

21. $\iint_R (x^2 + y^2)^2\,dx\,dy$, donde R es la región limitada por el círculo con centro en $(0, 4)$ y radio 4.

22. $\iint_R xy\,dx\,dy$, donde R es la región en el primer cuadrante, limitada por el círculo con centro en $(0, 1)$ y radio 1, y la recta $y = \sqrt{3}x$.

23. $\iint_R xy\,dx\,dy$, donde $R = \{(x, y)|x^2 + \frac{y^2}{16} \leq 1\}$.

24. $\iint_R (x - y)\,dx\,dy$, donde R es la región limitada la elipse $2x^2 + 3y^2 = 1$.

25. $\iint_R \frac{1}{4 - 3x^2 - y^2}\,dx\,dy$, donde $R = \{(x, y)|3x^2 + y^2 \leq 1\}$.

6.5 Aplicaciones de las integrales dobles

En esta sección presentamos varios ejemplos de aplicación de las integrales dobles en problemas de geometría y de mecánica.

6.5.1 Volúmenes de cuerpos en el espacio

Si la función $z = f(x, y)$ es continua y no negativa, se ha visto que la integral doble

$$\iint_R f(x, y)\, dx\, dy$$

nos da el volumen bajo la superficie de la gráfica de la función $f(x, y)$ sobre la región $R \subset \mathbb{R}^2$. Más aún, si $g(x, y)$ es otra función tal que $g(x, y) \leq f(x, y)$ para toda (x, y) en R, la integral

$$\iint_R (f(x, y) - g(x, y))\, dx\, dy$$

nos da el volumen atrapado entre las dos superficies $z = f(x, y)$ y $z = g(x, y)$ sobre la región R (sin importar si este volumen queda por encima o por debajo del plano xy).

Ejemplo 1. Se quiere calcular el volumen del cuerpo limitado por los planos $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, $x = 0$, $y = 0$, $z = 0$, donde a, b, c son 3 números positivos dados. Obsérvese que se trata de un tetraedro, una de cuyas caras está en el plano $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, y las otras tres están en los planos coordenados. En forma esquemática el cuerpo se ve como

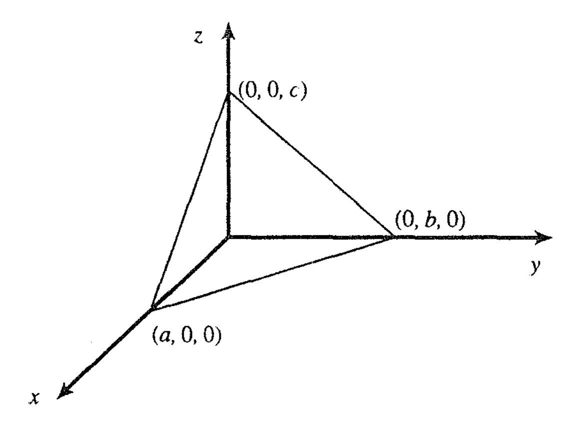


Figura 1. Cuerpo del ejemplo 1.

Despejando z de la ecuación $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, obtenemos la función $z = f(x, y) = c\left(1 - \frac{x}{a} - \frac{y}{b}\right)$, de la cual queremos obtener el volumen bajo su gráfica sobre la región que este tetraedro proyecta sobre el plano xy. Esta región es un triángulo con vértices en $(0, 0)$, $(a, 0)$ y $(0, b)$, y puede quedar descrita como una región del tipo I como

$$R = \{(x, y) | 0 \leq x \leq a, 0 \leq y \leq b(1 - x/a)\}$$

Así pues, el volumen procurado es

$$\begin{aligned}
V &= \iint_R c\left(1 - \frac{x}{a} - \frac{y}{b}\right) dx\, dy = \int_0^a \int_0^{b(1-x/a)} c\left(1 - \frac{x}{a} - \frac{y}{b}\right) dy dx \\
&= c\int_0^a dx \int_0^{b(1-x/a)} dy - \frac{c}{a}\int_0^a x\, dx \int_0^{b(1-x/a)} dy - \frac{c}{b}\int_0^a dx \int_0^{b(1-x/a)} y\, dy \\
&= c\int_0^a b(1 - x/a)\, dx - \frac{c}{a}\int_0^a b(1 - x/a)x\, dx - \frac{c}{2b}\int_0^a b^2(1 - x/a)^2\, dx \\
&= cb\left(x - \frac{x^2}{2a}\right)_0^a - \frac{cb}{a}\left(\frac{x^2}{2} - \frac{x^3}{3a}\right)_0^a + \frac{cba}{2}\left(\frac{(1 - x/a)^3}{3}\right)_0^a \\
&= cb\left(\frac{a}{2}\right) - \frac{cb}{a}\left(\frac{a^2}{6}\right) + \frac{cba}{2}\left(\frac{-1}{3}\right) = \frac{abc}{6}
\end{aligned}$$

Nótese que el resultado se puede escribir como $V = \frac{1}{3}\left(\frac{ab}{2}\right)c$, que corresponde a la fórmula que conocemos del volumen de un tetraedro como un tercio del área de su base por su altura. ■

Ejemplo 2. Se quiere calcular el volumen del cuerpo limitado por el cilindro[1] $z = 5 - 2x^2$, los planos coordenados y el plano $2x + y = 1$. Obsérvese que se trata de calcular el volumen bajo la superficie $z = f(x, y) = 5 - 2x^2$ sobre el triángulo formado por los ejes x e y y la recta $2x + y = 1$ (ahora nos referimos a esta ecuación como *una recta* pues estamos "hablando en $\mathbb{R}^2$"; al comienzo del problema nos referimos a esta misma ecuación como *un plano*, pues estábamos "hablando en $\mathbb{R}^3$"). Así, la región de integración puede ser vista como la región del tipo II siguiente

$$R = \{(x, y)|0 \leq y \leq 1, 0 \leq x \leq (1 - y)/2\}$$

de modo que el volumen procurado se calcula como

$$\begin{aligned}
\iint_R f(x, y)\, dx\, dy &= \int_0^1 \int_0^{(1-y)/2} (5 - 2x^2)\, dx\, dy \\
&= \int_0^1 \left(5x - \frac{2}{3}x^3\right)_0^{(1-y)/2} dy = \int_0^1 \left(\frac{5}{2}(1 - y) - \frac{1}{12}(1 - y)^3\right) dy \\
&= \left(-\frac{5}{4}(1 - y)^2 + \frac{1}{48}(1 - y)^4\right)_0^1 = \frac{5}{4} - \frac{1}{48} = \frac{59}{48}
\end{aligned}$$

(este es el número de unidades cúbicas que da el volumen procurado). ■

Ejemplo 3. Calculemos el volumen del cuerpo limitado por el paraboloide hiperbólico $z = xy$, el cilindro $y = \sqrt{2x}$ y los planos $x + y = 4$, $y = 0$, $z = 0$. Al imaginar el volumen que se pide calcular, se ve que la función que debemos integrar es la del paraboloide hiperbólico $z = xy$, y que

[1] La palabra *cilindro* se entiende en un sentido amplio: es una superficie que se extiende infinitamente de manera paralela a algún eje dado. Por ejemplo, $x^2 + y^2 = 1$ es un cilindro (en este caso, *circular recto*) cuya sección transversal es circular y se extiende paralelamente al eje z. De modo análogo $2y^2 + z^2 = 1$ es un cilindro elíptico recto que se extiende en forma paralela al eje x. En nuestro ejemplo, $z = 5 - 2x^2$ es un cilindro parabólico que se extiende en paralelo al eje y.

la región de integración está limitada por la parábola $y = \sqrt{2x}$ y las rectas $x + y = 4$, $y = 0$ (el eje x). El hecho de que $z = 0$ sea también "una pared" (de hecho es "el piso") del cuerpo cuyo volumen queremos calcular, nos dice simplemente que lo que vamos a calcular es un "volumen bajo la superficie $z = xy$". Al resolver simultáneamente las ecuaciones $y = \sqrt{2x}$ con $x + y = 4$ encontramos el punto de intersección de estas curvas en $\mathbf{p} = (2, 2)$. Así pues, la región de integración se describe fácilmente como una región del tipo II (si quisiéramos describirla como región del tipo I habría que dividir la región en 2 partes)

$$R = \{(x, y)|0 \leq y \leq 2, y^2/2 \leq x \leq 4 - y\}$$

como se muestra a continuación

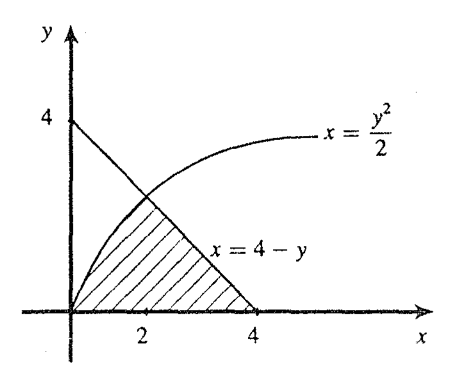


Figura 2. La región de integración del ejemplo 3.

Se tiene entonces que el volumen procurado es

$$V = \iint_R f(x, y)\,dx\,dy = \int_0^2 \int_{y^2/2}^{4-y} xy\,dx\,dy = \int_0^2 y\,dy\left(\frac{1}{2}x^2\right)_{y^2/2}^{4-y}$$

$$= \frac{1}{2}\int_0^2 y\left((4-y)^2 - y^4/4\right)dy$$

$$= \frac{1}{2}\left(8y^2 - \frac{8}{3}y^3 + \frac{1}{4}y^4 - \frac{1}{24}y^6\right)_0^2 = 6$$

■

Ejemplo 4. Calculemos el volumen debajo del paraboloide $f(x, y) = x^2 + y^2$ sobre el anillo circular $1 \leq x^2 + y^2 \leq 9$. Es claro que se trata de calcular la integral

$$V = \iint_R (x^2 + y^2)\,dx\,dy$$

sobre el anillo circular dado. Esta es una integral que se calcula mucho más fácilmente en coordenadas polares, en cuyo caso el integrando es la función $f(F(r, \theta)) = r^2$ y la región de integración es el rectángulo

$$R' = \{(r, \theta)|1 \leq r \leq 3, 0 \leq \theta \leq 2\pi\}$$

de modo que

$$V = \int_1^3 \int_0^{2\pi} r^2(r)\, dr d\theta = 2\pi \int_1^3 r^3\, dr = \frac{\pi}{2} r^4 \Big]_1^3 = 40\pi$$ ■

Ejemplo 5. Se quiere calcular el volumen del cuerpo comprendido entre la semiesfera $z = \sqrt{2 - x^2 - y^2}$ y el cono $z = \sqrt{x^2 + y^2}$ (se entiende la parte de la semiesfera que queda dentro del cono). La intersección de estas dos superficies se obtiene al igualar sus ecuaciones

$$\sqrt{2 - x^2 - y^2} = \sqrt{x^2 + y^2}$$

de donde $x^2 + y^2 = 1$. Se entiende entonces que debemos integrar la diferencia de la función $\sqrt{2 - x^2 - y^2}$ con la función $\sqrt{x^2 + y^2}$, sobre la región R limitada por el círculo unitario. Es decir, el volumen procurado es

$$V = \iint_R \left(\sqrt{2 - x^2 - y^2} - \sqrt{x^2 + y^2}\right) dx\, dy$$

Nuevamente, es claro que este cálculo resulta mucho más económico en coordenadas polares, quedándonos como

$$\begin{aligned} V &= \int_0^1 \int_0^{2\pi} \left(\sqrt{2 - r^2} - r\right) r\, dr d\theta \\ &= 2\pi \int_0^1 \sqrt{2 - r^2}\, r\, dr - 2\pi \int_0^1 r^2\, dr = -\frac{2\pi}{3}\left((2 - r^2)^{3/2} + r^3\right)_0^1 \\ &= -\frac{2\pi}{3}\left(2 - 2^{3/2}\right) = \frac{4\pi}{3}\left(\sqrt{2} - 1\right) \end{aligned}$$ ■

Ejemplo 6. Se quiere calcular el volumen que encierra el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Para efectuar tal cálculo, podemos despejar la z y establecer la función

$$z = f(x, y) = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}$$

cuya gráfica corresponde al semielipsoide superior ($z \geq 0$), y cuya proyección en el plano xy es la región limitada por la elipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

Entonces, el volumen procurado es dos veces el volumen del semielipsoide anterior. Es decir

$$V = 2 \iint_R c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\, dx\, dy = 2c \iint_R \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\, dx\, dy$$

donde $R = \left\{(x, y)|\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}$. Esta integral fue resuelta en el ejemplo 9 de la sección anterior, haciendo el cambio de variables a coordenadas polares generalizadas. Su valor es $\frac{2}{3}\pi ab$. Entonces el volumen del elipsoide es

$$V = 2c\left(\frac{2}{3}\pi ab\right) = \frac{4}{3}\pi abc$$

(obsérvese que como caso particular se puede obtener el volumen de la esfera de radio r, poniendo $a = b = c = r$, quedando $V = \frac{4}{3}\pi r^3$). ■

6.5.2 Áreas de figuras planas

Consideremos la función $z = f(x, y) = 1$. Al integrar esta función, que es continua y positiva, sobre una región R del plano xy, obtenemos el volumen del cilindro cuya base es R y cuya altura es 1. Al ser este volumen igual al área de su base (el área de R) por su altura (que es 1), vemos que lo que obtenemos del cálculo de la integral es numéricamente igual al área de la región de integración R. Así pues, se tiene la siguiente igualdad (numérica)

$$\iint_R dx\,dy = \text{ área de } R$$

Por ejemplo, si tenemos la siguiente región R del tipo I

$$R = \{(x, y)|a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$

obtendríamos que

$$\text{área de } R = \iint_R dx\,dy = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} dx\,dy = \int_a^b \big(\phi_1(x) - \phi_2(x)\big)dx$$

resultado que coincide con lo que ya sabíamos de nuestro primer curso de cálculo (es el área atrapada entre las gráficas de las funciones continuas $y = \phi_1(x)$, y, $y = \phi_2(x)$, calculada en el intervalo $[a, b]$). Por supuesto que la ventaja de calcular las áreas de regiones planas por medio de integrales dobles se hace palpable cuando se calcula dicha integral haciendo algún cambio conveniente de variables (cuando así lo amerite la situación).

Ejemplo 7. Consideremos las cuatro rectas

$$\ell_1: y = \alpha x + a, \quad \ell_2: y = \alpha x + b, \quad \ell_3: y = \beta x + c, \quad \ell_4: y = \beta x + d$$

en que $\alpha > \beta$, $b > a$, $d > c$. Se quiere calcular el área del paralelogramo con lados en cada una de estas rectas. Este es un problema geométrico que se puede se resolver de diversas maneras "elementales", pero, por supuesto, lo planteamos aquí para resolverlo con la herramienta que estamos estudiando. Podemos, por ejemplo, hallar los puntos de intersección de las cuatro rectas (los vértices del paralelogramo) y dividir la región de integración en tres subregiones del tipo I, con lo cual el área buscada se calcularía como

$$A = \int_{x_1}^{x_2} \int_{y_4}^{y_1} dx\,dy + \int_{x_2}^{x_3} \int_{y_2}^{y_1} dx\,dy + \int_{x_3}^{x_4} \int_{y_2}^{y_3} dx\,dy$$

en que y_i corresponde a la recta ℓ_i, $i = 1, 2, 3, 4$, y

$$x_1 = \frac{c-a}{\alpha-\beta}, \quad x_2 = \frac{c-b}{\alpha-\beta}, \quad x_3 = \frac{d-a}{\alpha-\beta}, \quad x_4 = \frac{d-b}{\alpha-\beta}$$

Sin embargo, el cálculo de esta área se podría hacer mucho más simple, si hacemos el cambio de variables según las fórmulas

$$u = y - \alpha x, \quad v = y - \beta x$$

En tal caso la región de integración corresponde al rectángulo

$$\{(u, v) | a \leq u \leq b, c \leq v \leq d\}$$

El jacobiano $\frac{\partial(u,v)}{\partial(x,y)}$ se calcula fácilmente

$$\frac{\partial(u,v)}{\partial(x,y)} = \det \begin{bmatrix} -\alpha & 1 \\ -\beta & 1 \end{bmatrix} = -(\alpha - \beta)$$

de modo que

$$\left| \frac{\partial(x,y)}{\partial(u,v)} \right| = \frac{1}{\alpha - \beta}$$

y entonces el área buscada es

$$A = \int_a^b \int_c^d \frac{1}{\alpha - \beta} du dv = \frac{(b-a)(d-c)}{\alpha - \beta}$$ ■

Ejemplo 8. La curva cuya ecuación es

$$(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$$

se conoce como "lemniscata de Bernoulli". Obsérvese que se trata de una curva simétrica respecto de los dos ejes coordenados. Su aspecto gráfico es el de la figura 3

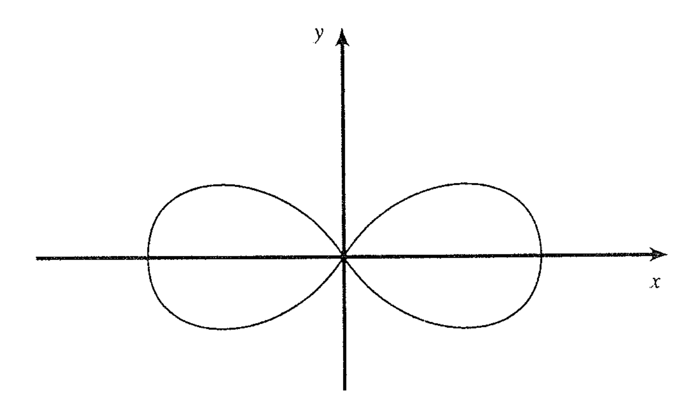


Figura 3. La lemniscata de Bernoulli

Queremos calcular el área que esta curva encierra. Por la simetría de la curva, basta calcular el área de la parte de la región encerrada por la curva en el primer cuadrante. Haciendo el cambio a coordenadas polares, $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, la ecuación de la curva queda como

$$(r^2\cos^2\theta + r^2\operatorname{sen}^2\theta)^2 = 2a^2(r^2\cos^2\theta - r^2\operatorname{sen}^2\theta)$$

o bien, simplificando

$$r = a\sqrt{2\cos 2\theta}$$

Nótese que en este caso hay restricciones acerca del dominio de esta función (valores de θ para los cuales existe r). La función está definida en función de los ángulos θ, $0 \leq \theta \leq 2\pi$, para los que $\cos 2\theta \geq 0$, corresponden a los intervalos $[0, \pi/4]$, $[3\pi/4, 5\pi/4]$, $[7\pi/4, 2\pi]$. De hecho, como calcularemos el área solamente en el primer cuadrante, consideraremos la variación de θ en el intervalo $[0, \pi/4]$. Así pues, el área procurada es

$$A = 4\int_0^{\pi/4}\int_0^{a\sqrt{2\cos 2\theta}} r\,dr d\theta = 2a^2\int_0^{\pi/4}(2\cos 2\theta)\,d\theta$$

$$= 2a^2\operatorname{sen} 2\theta\Big]_0^{\pi/4} = 2a^2$$ ■

Ejemplo 9. En el ejemplo 11 de la sección anterior, calculamos el área de la región encerrada por la curva

$$\left(\frac{x}{a} + \frac{y}{b}\right)^4 = \frac{x^2}{h^2} + \frac{y^2}{k^2}$$

en el primer cuadrante (a, b, h, k son números positivos dados), pues integramos la función $f(x, y) = 1$ sobre dicha región. Por lo que, dicha área es

$$A = \frac{ab}{6}\left(\frac{a^2}{h^2} + \frac{b^2}{k^2}\right)$$ ■

6.5.3 Centros de masa y momentos de figuras planas

Las integrales dobles se usan para calcular algunas magnitudes físicas de cuerpos que ocupan una región R en el plano. Comenzamos por introducir el concepto de "momento estático" de un punto de masa m respecto de un eje ℓ dado. Este momento, que denotaremos por M_ℓ, se define como el producto de la masa m del cuerpo por su distancia d al eje ℓ. Es decir, $M_\ell = md$. Si tenemos k cuerpos (puntuales) en el plano de masas m_i cuya distancia al eje ℓ es d_i, $i = 1, 2, \ldots, k$, entonces el momento estático de este sistema de cuerpos respecto del eje ℓ es

$$M_\ell = m_1d_1 + m_2d_2 + \cdots + m_kd_k = \sum_{i=1}^{k} m_id_i$$

Supongamos que el i-ésimo de estos cuerpos se encuentra en el punto $P_i = (x_i, y_i)$ del plano. Entonces los momentos estáticos respecto del eje x y respecto del eje y son

$$M_x = \sum_{i=1}^{k} m_iy_i, \quad M_y = \sum_{i=1}^{k} m_ix_i$$

El *centro de masa* del sistema es un punto $(\bar{x}, \bar{y})$ tal que si en él se concentrara toda la masa del sistema, los momentos estáticos de este punto respecto de los ejes coordenados son los mismos que los correspondientes de todo el sistema. Se debe tener entonces que

$$(\text{masa total del sistema})\ (\bar{y}) = M_x = \sum_{i=1}^{k} m_i y_i$$

$$(\text{masa total del sistema})\ (\bar{x}) = M_y = \sum_{i=1}^{k} m_i x_i$$

de donde

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i x_i}{\sum_{i=1}^{k} m_i}, \quad \bar{y} = \frac{\sum_{i=1}^{k} m_i y_i}{\sum_{i=1}^{k} m_i}$$

Si en lugar de considerar un sistema *discreto*, como el anterior, consideramos un sistema continuo en el que un cuerpo plano ocupa una región R del plano xy, el centro de masa del sistema se calcula con fórmulas similares donde "las sumatorias son sustituidas por integrales". Más aún: empecemos por establecer cómo se calcula la masa total de un cuerpo plano que ocupa una región R de $\mathbb{R}^2$ (diremos: "el cuerpo R"). Suponemos que, en general, el cuerpo *no es homogéneo*. Esto significa que hay una *función densidad* $\rho: R \to \mathbb{R}$ tal que a cada punto $(x, y) \in R$ asocia el valor de la densidad $\rho(x, y)$ del cuerpo en ese punto (en unidades de masa por unidades de área, por ejemplo gr/cm^2). La *masa total del cuerpo* R es entonces

$$M = \iint_R \rho(x, y)\, dx\, dy$$

(de manera intuitiva se puede pensar -como lo hacen los físicos- que si se multiplica la densidad $\rho(x, y)$ por el "elemento de área $dx\,dy$", se obtiene la masa puntual del cuerpo en el punto (x, y). En forma dimensional resulta aceptable este razonamiento, pues tomando a las coordenadas x, y en cm y pensando en dx y dy como "medidas infinitamente pequeñas de las coordenadas del punto (x, y)", se tendría que el producto de $\rho(x, y)$ en gr/cm^2 por $dx\,dy$ en cm^2 es efectivamente la cantidad $\rho(x, y)\,dx\,dy$ dada en gr. Para obtener la masa total del cuerpo R hay que *sumar de manera continua* todas estas masas puntuales sobre la región R; esto es lo que significa hacer la integral doble de $\rho(x, y)\,dxdy$ sobre R).

No será difícil aceptar entonces que los momentos estáticos del cuerpo R respecto de los ejes coordenados son

$$M_x = \iint_R y\rho(x, y)\, dx\, dy, \quad M_y = \iint_R x\rho(x, y)\, dx\, dy$$

y que las coordenadas del centro de masa del cuerpo son

$$\bar{x} = \frac{\iint_R x\rho(x, y)\, dx\, dy}{\iint_R \rho(x, y)\, dx\, dy}, \quad \bar{y} = \frac{\iint_R y\rho(x, y)\, dx\, dy}{\iint_R \rho(x, y)\, dx\, dy}$$

Nótese que si el cuerpo es homogéneo, lo cual significa que tiene densidad constante, digamos $\rho(x, y) = \rho_0$, entonces su masa total es

$$M = \iint_R \rho(x, y)\, dx\, dy = \rho_0 \iint_R dx\, dy = \text{(densidad)(área del cuerpo)}$$

y las coordenadas de su centro de masa se escribirían como

$$\bar{x} = \frac{\displaystyle\iint_R x\, dx\, dy}{\displaystyle\iint_R dx\, dy}, \quad \bar{y} = \frac{\displaystyle\iint_R y\, dx\, dy}{\displaystyle\iint_R dx\, dy}$$

Ejemplo 10. Se quiere calcular el momento estático de un círculo homogéneo (de densidad ρ_0) de radio r_0 respecto de una recta tangente a él. Para usar las fórmulas establecidas anteriormente de los momentos estáticos respecto de los ejes coordenados, tomemos el círculo tangente a alguno de estos ejes. Digamos, por ejemplo, que nuestro círculo es $x^2 + y^2 = 2r_0x$ (o bien $(x - r_0)^2 + y^2 = r_0^2$, el cual es un círculo con centro en $(r_0, 0)$ y radio r_0 y es, por tanto, tangente al eje y) y calculemos el momento estático M_y. Se tiene

$$M_y = \iint_R x\rho(x, y)\, dx\, dy = \rho_0 \iint_R x\, dx\, dy$$

Calculemos esta integral en coordenadas polares. En tal caso el círculo del problema se describe como $r = 2r_0 \cos\theta$, así que la variación en la coordenada r será de $r = 0$ a $r = 2r_0\cos\theta$, y la variación de θ será de $-\pi/2$ a $\pi/2$ (para describir así todo el círculo que se encuentra en el cuarto y primer cuadrante). Se tiene entonces

$$\begin{aligned}
M_y &= \rho_0 \iint_R x\, dx\, dy = \rho_0 \int_{-\pi/2}^{\pi/2} \int_0^{2r_0\cos\theta} (r\cos\theta) r\, dr d\theta \\
&= \rho_0 \int_{-\pi/2}^{\pi/2} \cos\theta \left(\frac{1}{3}r^3\right)_0^{2r_0\cos\theta} d\theta = \frac{8}{3}r_0^3\rho_0 \int_{-\pi/2}^{\pi/2} \cos^4\theta\, d\theta \\
&= \frac{8}{3}r_0^3\rho_0 \int_{-\pi/2}^{\pi/2} \left(\frac{1}{2}(1 + \cos 2\theta)\right)^2 d\theta \\
&= \frac{2}{3}r_0^3\rho_0 \int_{-\pi/2}^{\pi/2} \left(1 + 2\cos 2\theta + \frac{1}{2}(1 + \cos 4\theta)\right) d\theta \\
&= \frac{2}{3}r_0^3\rho_0 \int_{-\pi/2}^{\pi/2} \left(\frac{3}{2} + 2\cos 2\theta + \frac{1}{2}\cos 4\theta\right) d\theta \\
&= \frac{2}{3}r_0^3\rho_0 \left(\frac{3}{2}\theta + \operatorname{sen} 2\theta + \frac{1}{8}\operatorname{sen} 4\theta\right)_{-\pi/2}^{\pi/2} = \pi r_0^3\rho_0
\end{aligned}$$

■

Ejemplo 11. Obtengamos el centro de masa de un cuerpo homogéneo en forma de semicírculo de radio r_0. Podemos colocar entonces el cuerpo en la región

$$R = \{(x, y)| -r_0 \leq x \leq r_0, 0 \leq y \leq \sqrt{r^2 - x^2}\}$$

El área del cuerpo (el denominador de las coordenadas del centro de masa), es $\pi r^2/2$. Calculemos las integrales en los numeradores de $\bar{x}$ y $\bar{y}$. Se tiene

$$\iint_R x\,dx\,dy \xrightarrow{\text{cambio a polares}} = \int_0^{\pi}\int_0^{r_0} (r\cos\theta) r\,dr\,d\theta$$

$$= \frac{r_0^3}{3}\int_0^{\pi}\cos\theta d\theta = \frac{r_0^3}{3}(0) = 0$$

de modo que $\bar{x} = 0$ (esta conclusión la hubiéramos podido obtener con consideraciones de simetría del cuerpo R). Por otra parte

$$\iint_R y\,dx\,dy \xrightarrow{\text{cambio a polares}} = \int_0^{\pi}\int_0^{r_0} (r\,\text{sen}\,\theta) r\,dr d\theta$$

$$= \frac{r_0^3}{3}\int_0^{\pi}\text{sen}\,\theta\,d\theta = \frac{r_0^3}{3}\left(-\cos\theta\right)_0^{\pi} = \frac{2r_0^3}{3}$$

así que

$$\bar{y} = \frac{2r_0^3/3}{\pi r_0^2/2} = \frac{4r_0}{3\pi}$$

■

Ejemplo 12. (Un teorema de Pappus). Consideremos la región del tipo I

$$R = \{(x, y)|a \leq x \leq b, g(x) \leq y \leq f(x)\}$$

donde $f, g: [a, b] \to \mathbb{R}$ son funciones continuas y $f(x) > g(x) \geq 0\ \forall x \in [a, b]$. Sea S el cuerpo de revolución obtenido al girar la región R en torno al eje x. De nuestro primer curso de cálculo, sabemos que el volumen $V(S)$ de este cuerpo es

$$V(S) = \pi\int_a^b (f^2(x) - g^2(x))\,dx$$

Supongamos ahora que tenemos un cuerpo homogéneo en la región R. La ordenada de su centro de masa es, como vimos anteriormente

$$\bar{y} = \frac{\iint_R y\,dx\,dy}{\iint_R dx\,dy} = \frac{1}{\text{área de } R}\iint_R y\,dx\,dy$$

$$= \frac{1}{\text{área de } R}\int_a^b\int_{g(x)}^{f(x)} y\,dx = \frac{1}{\text{área de } R}\int_a^b \frac{1}{2}y^2\Big]_{g(x)}^{f(x)} dx$$

$$= \frac{1}{2}\frac{1}{\text{área de } R}\int_a^b (f^2(x) - g^2(x))\,dx = \frac{1}{2}\frac{1}{\text{área de } R}\frac{V(S)}{\pi}$$

Es decir que

$$\bar{y} = \frac{V(S)}{2\pi(\text{área de } R)}$$

de modo que es posible calcular la ordenada del centro de masa del cuerpo homogéneo sobre la región R conociendo su área y el volumen del cuerpo de revolución engendrado al girar esta región en torno al eje x. Este resultado es conocido como Teorema de Pappus. Para ilustrar el uso de este resultado, retomemos el ejemplo anterior. Al calcular el centro de masa de un cuerpo homogéneo en forma de semicírculo de radio r_0, consideramos la región R del tipo I en la que se encuentra el cuerpo

$$R = \{(x, y) | -r_0 \leq x \leq r_0, 0 \leq y \leq \sqrt{r^2 - x^2}\}$$

Por consideraciones de simetría, sólo podemos calcular la coordenada $\bar{y}$, pues $\bar{x} = 0$. El área de esta región es la mitad del área de un círculo de radio r_0, es decir, área de $R = \frac{1}{2}\pi r_0^2$. Al girar esta región en torno al eje de las x obtenemos una esfera de radio r_0 cuyo volumen es $V(S) = \frac{4}{3}\pi r_0^3$. Aplicando entonces el Teorema de Pappus, obtenemos que la coordenada $\bar{y}$ del centro de masa del cuerpo es

$$\bar{y} = \frac{V(S)}{2\pi(\text{área de } R)} = \frac{\frac{4}{3}\pi r_0^3}{2\pi\left(\frac{1}{2}\pi r_0^2\right)} = \frac{4r_0}{3\pi}$$

resultado que obtuvimos en el ejemplo anterior. ■

Ejemplo 13. Hallemos el centro de masa de un cuerpo en forma de semicírculo de radio 1, no homogéneo, cuya densidad en el punto (x, y) es proporcional al cuadrado de la distancia de este punto al punto $(0, -1)$ y en el origen vale 1 gr/cm^2. Obsérvese que se trata de un cuerpo en la misma región R del ejemplo 11, sólo que ahora dicho cuerpo se considera no homogéneo. La función densidad es en este caso $\rho(x, y) = k(x^2 + (y + 1)^2)$, donde k es una constante de proporcionalidad que podemos calcular pues sabemos que $\rho(0, 0) = 1 = k$. Así que $\rho(x, y) = x^2 + (y + 1)^2 = 1 + x^2 + y^2 + 2y$. La masa total del cuerpo es

$$M = \iint_R \rho(x, y)\, dx\, dy \xrightarrow{\text{cambio a polares}} = \int_0^{\pi} \int_0^1 (1 + r^2 + 2r \operatorname{sen}\theta) r\, dr d\theta$$

$$= \int_0^{\pi} \left(\frac{1}{2}r^2 + \frac{1}{4}r^4 + \frac{2}{3}r^3 \operatorname{sen}\theta\right)\Big|_0^1 d\theta = \int_0^{\pi} \left(\frac{3}{4} + \frac{2}{3}\operatorname{sen}\theta\right) d\theta$$

$$= \left(\frac{3}{4}\theta - \frac{2}{3}\cos\theta\right)\Big|_0^{\pi} = \frac{3}{4}\pi + \frac{4}{3} \text{ gr}$$

Por consideraciones de simetría podemos concluir que la abscisa del centro de masa del cuerpo R será cero (el cuerpo es simétrico respecto del eje y y la función densidad presenta también simetría respecto de este eje). Obtengamos entonces el numerador de la ordenada $\bar{y}$

$$\iint_R y\rho(x, y)\,dx\,dy \xrightarrow{\text{cambio a polares}} = \int_0^{\pi}\int_0^1 r \operatorname{sen}\theta\left(1 + r^2 + 2r \operatorname{sen}\theta\right) r\,dr d\theta$$

$$= \int_0^{\pi}\left(\frac{1}{3}r^3 \operatorname{sen}\theta + \frac{1}{5}r^5 \operatorname{sen}\theta + \frac{1}{2}r^4 \operatorname{sen}^2\theta\right)_0^1 d\theta$$

$$= \int_0^{\pi}\left(\frac{8}{15}\operatorname{sen}\theta + \frac{1}{4} - \frac{1}{4}\cos 2\theta\right) d\theta$$

$$= -\frac{8}{15}\cos\theta + \frac{1}{4}\theta - \frac{1}{8}\operatorname{sen} 2\theta\Big]_0^{\pi} = \frac{1}{4}\pi + \frac{16}{15}$$

de modo que la ordenada del centro de masa del cuerpo está en

$$\bar{y} = \frac{\frac{1}{4}\pi + \frac{16}{15}}{\frac{3}{4}\pi + \frac{4}{3}}$$

Nótese que si el cuerpo fuera homogéneo, la ordenada de su centro de masa estaría en (según el ejemplo 11) $\bar{y} = \frac{4}{3\pi} \approx 0.4244$. En nuestro caso está desplazada hacia arriba (en $\bar{y} \approx 0.50198$), pues el cuerpo es más denso en su parte superior. ■

Si tenemos un cuerpo plano sobre la región R con función densidad $\rho(x, y)$, el *momento de inercia* de este cuerpo respecto del eje x (del eje y , resp.), denotado por I_x (I_y resp.) es

$$I_x = \iint_R y^2\rho(x, y)\,dx\,dy$$

$$\left(I_y = \iint_R x^2\rho(x, y)\,dx\,dy, \text{ respectivamente}\right)$$

A la suma de estos dos momentos se le llama *momento polar de inercia* y se denota por I_0. Este es entonces

$$I_0 = I_x + I_y = \iint_R (x^2 + y^2)\rho(x, y)\,dx\,dy$$

Al momento de inercia de un cuerpo se le considera como una medida de la oposición a girar de éste cuando actúa en él una fuerza de rotación.

Ejemplo 14. Calculemos el momento de inercia de un cuerpo homogéneo de densidad ρ_0, de forma cuadrada de lado a, respecto de un eje que pasa por uno de sus lados. Podemos colocar el cuerpo en la región

$$R = \{(x, y)|0 \le x \le a, 0 \le y \le a\}$$

y calcular el momento de inercia respecto del eje x. Este es

$$I_x = \iint_R y^2\rho_0\,dx\,dy = \rho_0\int_0^a\int_0^a y^2\,dx\,dy = \rho_0\int_0^a\left(\frac{1}{3}y^3\right)_0^a dx$$

$$= \rho_0\int_0^a \frac{1}{3}(a)^3\,dx = \frac{1}{3}\rho_0 a^3(a) = \frac{1}{3}\rho_0 a^4$$ ■

6.5.4 Valor medio de una función

Consideremos una función $f: R \subset \mathbb{R}^2 \to \mathbb{R}$ continua definida en una región R del plano xy. El *valor medio* de la función f sobre la región R, denotado por $\bar{f}_R$ se define por

$$\bar{f}_R = \frac{\iint_R f(x, y)\, dx\, dy}{\iint_R dx\, dy}$$

Nótese que el denominador de esta expresión es el área de la región R. El valor $\bar{f}_R$ es una medida del promedio de los valores de la función sobre la región R. Así, si $f(x, y)$ es, por ejemplo, una función de densidad (como las funciones $\rho(x, y)$ de la subsección anterior), $\bar{f}_R$ es una medida de la *densidad media* del cuerpo sobre la región R.

Podemos ser más precisos en el contenido de la fórmula para $\bar{f}_R$: el valor $\bar{f}_R$ lo toma la función f *en un punto* (a, b) *de la región* R. En realidad, este es un resultado análogo al Teorema del Valor Medio para Integrales, estudiado en nuestro primer curso de cálculo y que establecemos a continuación.

Teorema 6.5.1 (Teorema del valor medio para integrales dobles) Sea $f: R \subset \mathbb{R}^2 \to \mathbb{R}$ una función continua definida en la región R. Hay pues un punto $(a, b) \in R$ tal que

$$f(a, b) = \bar{f}_R = \frac{\iint_R f(x, y)\, dx\, dy}{\iint_R dx\, dy}$$ ■

La demostración de este resultado requiere algunos detalles técnicos sobre funciones continuas definidas en conjuntos compactos del plano. Aunque no la presentamos con todos los detalles, la idea general es la siguiente: la función f alcanza su máximo y mínimo absoluto (M y m respectivamente) en R, de modo que

$$m \leq f(x, y) \leq M$$

para toda (x, y) en R. Aplicando la propiedad **b.** del teorema 6.3.1, podemos escribir entonces que

$$\iint_R m\, dx\, dy \leq \iint_R f(x, y)\, dx\, dy \leq \iint_R M\, dx\, dy$$

de donde

$$m \leq \frac{\iint_R f(x, y)\, dx\, dy}{\iint_R dx\, dy} \leq M$$

El valor $\bar{f}_R$ que se encuentra en esta desigualdad es un valor entre el máximo y mínimo absolutos de la función f en R. Se puede demostrar entonces que éste es un valor que *toma* la función f en algún punto (a, b) de R, con algunas propiedades tanto de la continuidad de f como de las características topológicas de la región R. Con esto concluye el teorema. Si la función es no negativa, este resultado tiene una interpretación geométrica interesante (análoga a la de su símil para funciones de una variable): se da un cilindro de base R y altura $\bar{f}_R$ cuyo volumen (que es igual a $(\bar{f}_R)\iint_R dx\,dy$) es el mismo que el volumen bajo la superficie de $f(x, y)$ sobre R (que es $\iint_R f(x, y)\,dx\,dy$).

Ejemplo 14. El valor medio de una función constante $f(x, y) = c$ en cualquier región R es (obviamente) c. El valor medio de la función

$$f(x, y) = \sqrt{R^2 - x^2 - y^2}$$

sobre la región $T = \{(x, y)|x^2 + y^2 \leq R^2\}$ lo podemos calcular fácilmente con argumentos geométricos. El numerador de la fórmula que define a $\bar{f}_T$ es el volumen bajo la superficie de $f(x, y)$ sobre T. Este es la mitad del volumen de una esfera de radio R, es decir $\frac{2}{3}\pi R^3$. El denominador de tal fórmula es el área de la región T, o sea el área de un círculo de radio R, que es πR^2. Así pues, el valor medio de f sobre T es

$$\bar{f}_T = \frac{\frac{2}{3}\pi R^3}{\pi R^2} = \frac{2}{3}R$$ ■

Ejemplo 15. Consideremos el cuerpo no homogéneo del ejemplo 13, cuya función densidad es $\rho(x, y) = x^2 + (y + 1)^2$ sobre la región R de forma de semicírculo de radio 1. La densidad media de este cuerpo es

$$\bar{\rho} = \frac{\iint_R \rho(x, y)\,dx\,dy}{\iint_R dx\,dy} = \frac{2}{\pi}\iint_R \rho(x, y)\,dx\,dy$$

(puesto que el área de la región R es $\pi/2$). La integral cuyo cálculo falta es la masa total del cuerpo, que ya calculamos en el ejemplo 13, y que resultó ser $\frac{3}{4}\pi + \frac{4}{3}$. Así, el valor de la densidad media del cuerpo es

$$\bar{\rho} = \frac{2}{\pi}\left(\frac{3}{4}\pi + \frac{4}{3}\right) \approx 2.3488 \text{ gr/cm}^2$$ ■

Ejercicios (Capítulo 6, Sección 5)

En los ejercicios 1—10, hallar el volumen del cuerpo en $\mathbb{R}^3$ limitado por las superficies indicadas.

1. $z = x^2 + y^2, z = 1$.

2. $z = 5 - x - y, z = 0, x = 0, x = 2, y = 0, y = 3$.

3. $x + y + z = 6, x^2 + y^2 = 1, x = 0, y = 0, z = 0$.

4. $y = \ln x, y = \ln^2 x, x = 1, x = e, z = 0, z = 3$.

5. $z = xy, x + y = 2, z = 0.$
6. $z = \exp(-x^2 - y^2), x^2 + y^2 = 1.$
7. $x^2 + y^2 = 3, z = 0, x + y + z = 4.$
8. $z = (x^2 + y^2)^{1/2}, z = 2.$
9. $x^2 + y^2 + z^2 = 6, z = x^2 + y^2, z \geq 0.$
10. $z = x^2 + y^2, z = 8 - x^2 - y^2.$

En los ejercicios 11—15, calcule el volumen del cuerpo, limitado por las superficies dadas. Se recomienda hacer un cambio adecuado de variables en las integrales dobles que aparezcan en el problema.

11. $z = 2x^2 + y^2, z = 3.$
12. $z = x^2 + 4y^2 - 2, z = 2 - x^2 - 4y^2.$
13. $2x^2 + 3y^2 + z^2 = 6, z = (2x^2 + 3y^2)^{1/2}, z \geq 0.$
14. $z^2 = xy, x + y = 1, x + y = 2.$
15. $z = x^2 + y^2, xy = 1, xy = 2, 2y = x, 2x = y, z = 0.$

En los ejercicios 16—30, calcular el área de la región limitada por las curvas dadas.

16. $y^2 = x, x^2 = y.$
17. $y = x^2, y = x + 4.$
18. $y^2 = x, x = 5.$
19. $y = \ln x, y = -\ln x, x = e.$
20. $y = \ln x, y = \ln^2 x.$
21. $y = \operatorname{sen} x, y = \cos x, x = \pi/4, x = 5\pi/4.$
22. $y = \dfrac{1}{1 + x^2}, y = x^2/2.$
23. $(x^2 + y^2)^2 = x^3 - 3xy^2, x \geq 0, y \geq 0.$
24. $(x^2 + y^2)^2 = x^3 + y^3, x \geq 0, y \geq 0.$
25. $(x + y)^3 = xy, x \geq 0, y \geq 0.$
26. $x^2 + y^2 = x + y, x \geq 0, y \geq 0.$
27. $9x^2 + 4y^2 = 36(x + y)^4, x \geq 0, y \geq 0.$
28. $16x^2 + y^2 = 16(x^3 + y^3), x \geq 0, y \geq 0.$
29. $y - x = 1, y - x = -1, y + x = 2, y + x = -3.$
30. $xy = 2, xy = 4, y = x, y = 3x, x \geq 0, y \geq 0.$
31. Desde el punto $P = (0, \alpha)$ se han trazado tangentes a la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, (en donde $\alpha > b > 0$). Hallar el área del "triángulo" formado por las tangentes y la elipse.

32. Hallar el área de cada una de las tres partes en que la hipérbola $x^2 - 2y^2 = 1$ corta al círculo $x^2 + y^2 = 4$.

33. Hallar el área de cada una de las dos partes en que la parábola $y = 5x^2/8$ corta al círculo $x^2 + y^2 = 9/4$.

34. Calcule el área de la región comprendida entre los dos círculos $x^2 + y^2 = ax$, $x^2 + y^2 = by$, en donde a y b son dos números positivos dados.

35. Demuestre que la región limitada por los cuatro círculos tangentes entre sí $(x \pm \alpha)^2 + (y \pm \alpha)^2 = \alpha^2$ tiene área igual a $\alpha^2(4 - \pi)$, **a.** usando argumentos de geometría elemental, **b.** usando el resultado del ejercicio anterior.

36. Calcule el área de cada una de las partes en que la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ corta a la elipse $\frac{x^2}{b^2} + \frac{y^2}{a^2} = 1$, en donde $a > b > 0$.

37. Considere los dos círculos $x^2 + y^2 = ax$, $x^2 + y^2 = bx$, en donde $a > 0$, $b < 0$, $|b| < a$. Calcule el área del "triángulo" limitado por estos dos círculos y la recta tangente común a ellos.

38. Calcule el área del círculo $x^2 + y^2 = 2ay$, exterior al círculo $x^2 + y^2 = a^2$.

39. Los aros olímpicos pueden ser representados por las ecuaciones

$$\begin{aligned} x^2 + y^2 - 2y &= 0 \\ 16x^2 + 16y^2 - 40x + 9 &= 0 \\ 4x^2 + 4y^2 - 20x - 8y + 25 &= 0 \\ 16x^2 + 16y^2 - 120x + 209 &= 0 \\ x^2 + y^2 - 10x - 2y + 25 &= 0 \end{aligned}$$

Demuestre que la intersección entre cualesquiera dos de estos aros o es vacía o limita una región con un área aproximada de 0.326065 unidades cuadradas.

40. Demuestre que el área limitada por la curva (dada en coordenadas polares) $r = f(\theta)$, entre los ángulos $\theta = \theta_1$ y $\theta = \theta_2$, viene dada por

$$A = \frac{1}{2}\int_{\theta_1}^{\theta_2} f^2(\theta)\, d\theta$$

41. Calcular el área encerrada por la cardioide $r = a(1 + \cos\theta)$.

42. Calcular el área encerrada por el caracol de Pascal $r = 2 + \cos\theta$.

43. Calcular el área de cada uno de los pétalos de la rosa de cuatro pétalos $r = a|\operatorname{sen} 2\theta|$.

44. Hallar la masa total de una lámina cuadrada de lado α, si la densidad de la lámina en un punto P de ella es proporcional al cuadrado de la distancia entre P y el centro de la lámina, siendo en cada uno de sus vértices igual a k.

45. Hallar la masa total de una lámina en forma de triángulo equilátero de lado α, si la densidad en un punto P de ella es proporcional a la distancia entre P y uno de los vértices de la lámina, siendo en los dos vértices restantes igual a k.

En los ejercicios 46—50, hallar las coordenadas del centro de masa de la figura plana homogénea dada.

46. Triángulo isósceles de base b y altura h.

47. Sector circular de radio R y ángulo central θ.

48. Figura limitada por $y = x^2$, $y = 1$.

49. Figura limitada por $y = x^2$, $y = 5x - 6$.

50. Figura limitada por la línea cerrada $y^2 = x^2 - x^4$, $x \geq 0$.

En los ejercicios 51—55, calcular los momentos de inercia indicados.

51. De un rectángulo de base a y altura b, respecto del lado a.

52. De la región $R = \{(x, y)|x^2 + y^2 \leq r^2, y \geq 0\}$ respecto del eje x.

53. De la región $R = \{(x, y)|y \leq \operatorname{sen} x, 0 \leq x \leq \pi\}$ respecto del eje x.

54. De un círculo de radio r respecto de su tangente.

55. De un triángulo isósceles de base a y altura h, respecto de su base.

En los ejercicios 56—60, calcule el valor medio de la función $z = f(x, y)$ dada, en la región indicada.

56. $f(x, y) = xy$, en el cuadrado $R = [0, 1] \times [0, 1]$.

57. $f(x, y) = x^2 + y^2$, en la región $R = \{(x, y)|x + y \leq 1, x \geq 0, y \geq 0\}$.

58. $f(x, y) = c$, en una región R arbitraria.

59. $f(x, y) = x + y$, en el rectángulo $R = [0, 2] \times [0, 3]$.

60. $f(x, y) = e^{-(x^2+y^2)}$, en la región $R = \{(x, y)|x^2 + y^2 \leq 1\}$.

61. Determine el valor promedio del producto de dos números, si cada uno de estos varía entre 0 y 1. (Ver ejercicio 56).

62. Determine el valor promedio de la suma de los cuadrados de dos números no negativos, si estos varían de tal modo que su suma nunca es mayor a uno. (Ver ejercicio 57).

6.6 Integrales triples

En esta sección haremos una presentación general tanto de la parte teórica como de los métodos de cálculo para las integrales triples. Una vez desarrollada con cierto cuidado la teoría correspondiente a integrales dobles (como lo hemos hecho en las secciones anteriores), el paso a "una integral más con funciones de una variable más" resultará natural.

El primer paso que daremos será definir integral triple para funciones de tres variables $f\colon Q \subset \mathbb{R}^3 \to \mathbb{R}$ escalonadas definidas en productos cartesianos de 3 intervalos de $\mathbb{R}$, $Q = I_1 \times I_2 \times I_3$.

Sean $I_1 = [a, b]$, $I_2 = [c, d]$, $I_3 = [e, g]$ tres intervalos de $\mathbb{R}$. Llamaremos al producto cartesiano

$$Q = I_1 \times I_2 \times I_3 = \{(x, y, z)|x \in [a, b], y \in [c, d], z \in [e, g]\}$$

rectángulo en $\mathbb{R}^3$ (esto por vía de simplificación; un nombre más apropiado sería "paralelepípedo rectangular"). Una partición P de Q es un producto cartesiano del tipo $P = P_1 \times P_2 \times P_3$ en el que P_i es una partición del intervalo I_i, $i = 1, 2, 3$. Digamos que P_i divide al intervalo I_i en m_i subintervalos. Entonces P divide a Q en $m_1 m_2 m_3$ subrectángulos en $\mathbb{R}^3$ y escribiremos $Q_{ijk} = [x_i, x_{i-1}] \times [y_j, y_{j-1}] \times [z_k, z_{k-1}]$ para denotar al ijk-ésimo subrectángulo de la partición P de Q, $i = 1, 2, \ldots, m_1$, $j = 1, 2, \ldots, m_2$, $k = 1, 2, \ldots, m_3$.

Se dice que la función $f\colon Q \subset \mathbb{R}^3 \to \mathbb{R}$ definida en el rectángulo Q en $\mathbb{R}^3$ es una *función escalonada*, si existe una partición P de Q, digamos que en $m_1 m_2 m_3$ subrectángulos Q_{ijk}, de modo que f restringida es constante a cada uno de estos subrectángulos Q_{ijk}. Digamos que $f\big|_{Q_{ijk}} = c_{ijk} \in \mathbb{R}$. Se define la *integral triple* de la función f sobre Q, denotada por

$$\iiint\limits_Q f \quad \text{o} \quad \iiint\limits_Q f(x, y, z)\, dx\, dy\, dz$$

como

$$\iiint\limits_Q f(x, y, z)\, dx\, dy\, dz = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sum_{k=1}^{m_3} c_{ijk}(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})$$

Nótese que el producto $(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})$ es el volumen del ijk-ésimo subrectángulo $Q_{ijk} = [x_i, x_{i-1}] \times [y_j, y_{j-1}] \times [z_k, z_{k-1}]$ de la partición P de Q, así que la integral triple de f en Q la se puede escribir como

$$\iiint\limits_Q f(x, y, z)\, dx\, dy\, dz = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sum_{k=1}^{m_3} \big(\text{ valor de } f \text{ en} Q_{ijk}\big)\big(\text{ volumen de } Q_{ijk}\big)$$

Una vez establecido el concepto de integral triple para funciones escalonadas, éste se establece ahora para funciones acotadas (como se hizo en la sección 2).

Definición. Sea $f\colon Q \subset \mathbb{R}^3 \to \mathbb{R}$ una función acotada definida en el rectángulo Q de $\mathbb{R}^3$. Si hay un único número real I tal que

$$\iiint\limits_Q \phi(x, y, z)\, dx\, dy\, dz \le I \le \iiint\limits_Q \psi(x, y, z)\, dx\, dy\, dz$$

para todas las funciones escalonadas $\phi, \psi\colon Q \subset \mathbb{R}^3 \to \mathbb{R}$ de modo que

$$\phi(x, y, z) \le f(x, y, z) \le \psi(x, y, z) \quad \forall (x, y, z) \in Q$$

se dice que la función f es *integrable* sobre Q, y al número I se le llama integral triple de f sobre Q y se escribe

$$I = \iiint\limits_Q f(x, y, z)\, dx\, dy\, dz$$

■

Hay un análogo al teorema 6.2.1 para integrales triples. Dejamos al lector que, a manera de ejercicio, enuncie y demuestre este resultado. Por supuesto que también vale un análogo al teorema 6.2.2 (funciones continuas sobre rectángulos son integrables), el cual en este caso establece que

Teorema 6.6.1 Sea $f: Q \subset \mathbb{R}^3 \to \mathbb{R}$ una función continua definida en el rectángulo $Q = [a, b] \times [c, d] \times [e, g]$ de $\mathbb{R}^3$. Entonces f es integrable en Q. ■

Podemos ser más precisos en la conclusión del teorema anterior e indicar como se puede calcular la integral de f sobre Q. Se tiene de hecho que

$$\iiint_Q f(x, y, z)\, dx\, dy\, dz = \int_a^b \int_c^d \left(\int_e^g f(x, y, z)\, dz \right) dx\, dy$$

En la expresión anterior, la integral que aparece en el paréntesis se realiza respecto de la variable z, quedando el resultado como una función de x y y, digamos $\xi(x, y)$. Entonces la integral triple de f en Q es la integral doble de $\xi(x, y)$ sobre el rectángulo de $\mathbb{R}^2$ $[a, b] \times [c, d]$. Más aún, podemos intercambiar el orden de integración y escribir que

$$\iiint_Q f(x, y, z)\, dx\, dy\, dz = \int_a^b \int_e^g \left(\int_c^d f(x, y, z)\, dy \right) dx\, dz$$

donde ahora la primera integral se realiza respecto de la variable y, quedando el resultado en función de x y z, digamos $\zeta(x, z)$. Así que la integral triple de f sobre Q es la integral doble de $\zeta(x, z)$ en el rectángulo $[a, b] \times [e, g]$. De manera análoga se tiene que

$$\iiint_Q f(x, y, z)\, dx\, dy\, dz = \int_c^d \int_e^g \left(\int_a^b f(x, y, z)\, dx \right) dy\, dz$$

El análogo al teorema 6.2.4 para integrales triples es

Teorema 6.6.2 Sea $f: Q \subset \mathbb{R}^3 \to \mathbb{R}$ una función definida y acotada en el rectángulo Q de $\mathbb{R}^3$. Si el conjunto de discontinuidades de f en Q tiene medida cero, entonces f es integrable en Q. ■

Hay que aclarar lo que significa que un conjunto en $\mathbb{R}^3$ tenga medida cero. Diremos que el conjunto $C \subset \mathbb{R}^3$ tiene medida cero si se dan rectángulos Q_i de $\mathbb{R}^3$, $i = 1, 2, \ldots, n$ cuya unión contiene a C y cuya suma de volúmenes es menor que cualquier número $\epsilon > 0$ previamente dado. Un ejemplo importante de conjunto de medida cero en $\mathbb{R}^3$ es el dado por la gráfica de una función continua $\phi: K \subset \mathbb{R}^2 \to \mathbb{R}$ definida en un conjunto compacto (cerrado y acotado) de $\mathbb{R}^2$ (como sabemos, esta es una superficie en $\mathbb{R}^3$). Este ejemplo es el análogo del teorema 6.2.3.

Para extender el concepto de integral triple de funciones continuas de tres variables a regiones más generales en $\mathbb{R}^3$, que se denota por Ω (en contraste con la letra R que usamos para regiones en $\mathbb{R}^2$), aclaramos en principio los tipos de regiones que se pueden presentar: en general éstas serán subconjuntos de $\mathbb{R}^3$ limitados por gráficas de funciones de dos variables. Por ejemplo, un tipo de

regiones que se van a considerar son las que están limitadas *por encima* y *por debajo* por gráficas de funciones de las variables x, y. De modo más preciso, si R es una región en el plano x, y (de las que estudiamos en la sección 3), definamos Ω como

$$\Omega = \{(x, y, z)|(x, y) \in R, \phi_1(x, y) \leq z \leq \phi_2(x, y)\}$$

donde $\phi_1, \phi_2: R \subset \mathbb{R}^2 \to \mathbb{R}$ son funciones continuas definidas en la región R de $\mathbb{R}^2$. Esquema de Ω

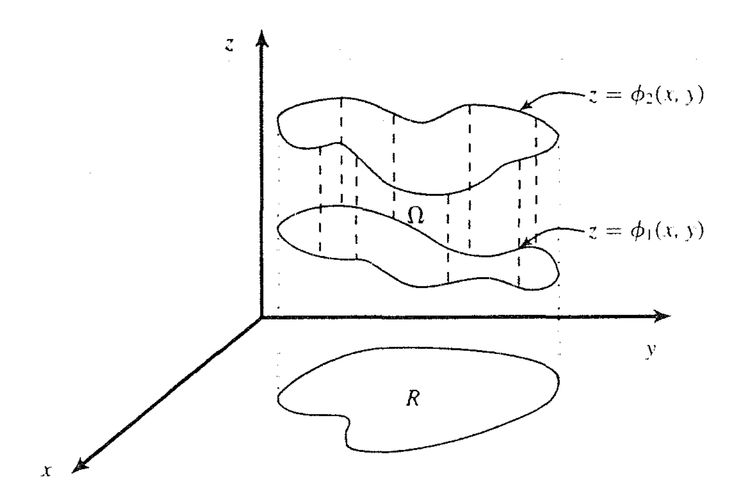


Figura 1. La región de Ω de $\mathbb{R}^3$

De hecho este tipo de regiones en $\mathbb{R}^3$ son acotadas. Es decir, se puede conseguir un rectángulo Q en $\mathbb{R}^3$ que contenga la región. Para introducir el concepto de integral triple de una función continua $f(x, y, z)$ sobre la región Ω, tomamos el rectángulo Q en $\mathbb{R}^3$ que contiene a Ω y definimos en Q la función $\tilde{f}(x, y, z)$ como

$$\tilde{f}(x, y, z) = \begin{cases} f(x, y, z) & \text{si } (x, y, z) \in \Omega \\ 0 & \text{si } (x, y, z) \in Q \backslash \Omega \end{cases}$$

Se puede demostrar que, siendo la función f continua en Ω, la función $\tilde{f}(x, y, z)$ es integrable sobre Q (sus discontinuidades forman un conjunto de medida cero, en el que están incluídos los puntos de las gráficas de las funciones $z = \phi_1(x, y)$ y $z = \phi_2(x, y)$). Se define entonces la integral de $f(x, y, z)$ sobre Ω como la integral de $\tilde{f}(x, y, z)$ sobre Q. Es decir

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iiint\limits_{Q} \tilde{f}(x, y, z)\, dx\, dy\, dz$$

Más aún, se puede ver que esta integral se calcula como

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iint\limits_{R} \left[\int_{\phi_1(x, y)}^{\phi_2(x, y)} f(x, y, z)\, dz\right] dx\, dy$$

La integral que aparece en el corchete se calcula respecto de la variable z, considerando constantes a x y y constantes, de modo que al bajar los límites de integración (en la variable z), los cuales son funciones de x y y, el resultado quedará como una función de las variables x y y, digamos $\xi(x, y)$. Entonces la integral triple de $f(x, y, z)$ sobre Ω es la integral doble de la función $\xi(x, y)$ sobre la región R, la cual se puede ver como *la proyección de la región Ω sobre el plano xy.*

Los resultados sobre "intercambio en el orden de integración" se obtienen fácilmente observando que las letras x, y, z de la exposición anterior juegan un papel completamente intercambiable. Por ejemplo, podemos considerar regiones Ω en $\mathbb{R}^3$ que estén limitadas "por los lados" por gráficas de funciones del tipo $y = \psi(x, z)$ y cuya proyección sea una región R en el plano xz. En forma más precisa, si R es una región en el plano xz, se considera la región Ω de $\mathbb{R}^3$ dada por

$$\Omega = \{(x, y, z)|(x, z) \in R,\ \psi_1(x, z) \leq y \leq \psi_2(x, z)\}$$

en la que $\psi_1, \psi_2: R \subset \mathbb{R}^2 \to \mathbb{R}$ son funciones continuas definidas en la región R del plano xz. En tal caso la integral triple de $f(x, y, z)$ sobre Ω se calcula como

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iint\limits_{R} \left[\int_{\psi_1(x,z)}^{\psi_2(x,z)} f(x, y, z)\, dy\right] dx\, dz$$

Por último, si tenemos una región R en el plano yz y consideramos la región Ω de $\mathbb{R}^3$ dada por

$$\Omega = \{(x, y, z)|(y, z) \in R,\ \upsilon_1(y, z) \leq x \leq \upsilon_2(y, z)\}$$

donde $\upsilon_1, \upsilon_2: R \subset \mathbb{R}^2 \to \mathbb{R}$ son funciones continuas definidas en la región R del plano yz. En tal caso la integral triple de $f(x, y, z)$ sobre Ω se calcula como

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iint\limits_{R} \left[\int_{\upsilon_1(y,z)}^{\upsilon_2(y,z)} f(x, y, z)\, dx\right] dy\, dz$$

Para finalizar este tema y antes de estudiar algunos ejemplos, presentamos el análogo al teorema 6.3.1 en el que se resumen las principales propiedades de las integrales triples de funciones continuas.

Teorema 6.6.3 Sean $f: \Omega \subset \mathbb{R}^3 \to \mathbb{R}$, $g: \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ dos funciones continuas definidas en la región Ω.

a. Si α y β son dos números reales cualesquiera, entonces

$$\iiint\limits_{\Omega} \left(\alpha f(x, y, z) + \beta g(x, y, z)\right) dx\, dy\, dz$$

$$= \alpha \iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz + \beta \iiint\limits_{\Omega} g(x, y, z)\, dx\, dy\, dz$$

b. Si $f(x, y, z) \geq g(x, y, z)$ para toda $(x, y, z) \in \Omega$, entonces

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz \geq \iiint\limits_{\Omega} g(x, y, z)\, dx\, dy\, dz$$

c. Si la región Ω está dividida en dos subregiones Ω_1 y Ω_2 (es decir, si se tiene $\Omega = \Omega_1 \cup \Omega_2$), entonces

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iiint\limits_{\Omega_1} f(x, y, z)\, dx\, dy\, dz + \iiint\limits_{\Omega_2} f(x, y, z)\, dx\, dy\, dz \qquad \blacksquare$$

El resto de la presente sección lo dedicaremos a presentar ejemplos resueltos de cálculo de integrales triples. Dejaremos la problemática relacionada con el cambio de variables para la siguiente sección.

Ejemplo 1. Sea $f(x, y, z)$ una función continua en $\mathbb{R}^3$. Se quiere calcular la integral triple de f en la región Ω de $\mathbb{R}^3$ limitada por los planos coordenados y el plano $x + y + z = 1$. Esta región es un tetraedro con vértices en los puntos $A = (0, 0, 0)$, $B = (1, 0, 0)$, $C = (0, 1, 0)$, $D = (0, 0, 1)$. La proyección de esta región sobre el plano xy es un triángulo con vértices en A, B y C, que, como región del plano se puede ver como

$$R = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$$

de modo que Ω se ve como

$$\Omega = \{(x, y, z) | (x, y) \in R, 0 \leq z \leq 1 - x - y\}$$

y entonces la integral triple de $f(x, y, z)$ sobre Ω es

$$\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz = \int_0^1 \left[\int_0^{1-x} \left(\int_0^{1-x-y} f(x, y, z)\, dz\right) dy\right] dx$$

Por ejemplo, si $f(x, y, z) = (x + 2y + 3z)^2$ se tiene

$$\begin{aligned}
\iiint\limits_{\Omega} f(x, y, z)\, dx\, dy\, dz &= \int_0^1 \left[\int_0^{1-x} \left(\int_0^{1-x-y} (x + 2y + 3z)^2\, dz\right) dy\right] dx \\
&= \int_0^1 \left[\int_0^{1-x} \left(\frac{1}{9}(x + 2y + 3z)^3\right)_0^{1-x-y} dy\right] dx \\
&= \frac{1}{9}\int_0^1 \left[\int_0^{1-x} \left((3 - 2x - y)^3 - (x + 2y)^3\right) dy\right] dx \\
&= \frac{1}{9}\int_0^1 \left[-\frac{1}{4}(3 - 2x - y)^4 - \frac{1}{8}(x + 2y)^4\right]_0^{1-x} dx \\
&= -\frac{1}{36}\int_0^1 \left((2 - x)^4 - (3 - 2x)^4\right) dx - \frac{1}{72}\int_0^1 \left((2 - x)^4 - x^4\right) dx \\
&= \left[\frac{1}{120}(2 - x)^5 - \frac{1}{360}(3 - 2x)^5 + \frac{1}{360}x^5\right]_0^1 = \frac{5}{12} \qquad \blacksquare
\end{aligned}$$

Ejemplo 2. Calculemos la integral triple

$$\iiint_{\Omega} \sqrt{R^2 - x^2 - y^2 - z^2}\, dx\, dy\, dz$$

donde Ω es la región limitada por la esfera $x^2 + y^2 + z^2 = R^2$. La proyección de esta región sobre el plano xy es

$$T = \{(x, y) | x^2 + y^2 \leq R^2\}$$

y está limitada por debajo por la semiesfera $z = -\sqrt{R^2 - x^2 - y^2}$ y por arriba por la semiesfera $z = \sqrt{R^2 - x^2 - y^2}$. Entonces

$$\iiint_{\Omega} \sqrt{R^2 - x^2 - y^2 - z^2}\, dx\, dy\, dz$$

$$= \int_{-R}^{R} \left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \left[\int_{-\sqrt{R^2-x^2-y^2}}^{\sqrt{R^2-x^2-y^2}} \sqrt{R^2 - x^2 - y^2 - z^2}\, dz \right] dy \right) dx$$

Con

$$\int \sqrt{a^2 - u^2}\, du = \frac{1}{2} u \sqrt{a^2 - u^2} + \frac{1}{2} a^2 \operatorname{arcsen} \frac{u}{a}$$

nos queda, haciendo la integral respecto de z y llamando $a = \sqrt{R^2 - x^2 - y^2}$

$$\int_{-R}^{R} \left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \left[\frac{1}{2} z \sqrt{a^2 - z^2} + \frac{1}{2} a^2 \operatorname{arcsen} \frac{z}{a} \right]_{-a}^{a} dy \right) dx$$

$$= \int_{-R}^{R} \left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{2} \pi a^2\, dy \right) dx$$

$$= \frac{1}{2} \pi \int_{-R}^{R} \left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} (R^2 - x^2 - y^2)\, dy \right) dx$$

$$= \frac{2}{3} \pi \int_{-R}^{R} (R^2 - x^2)^{3/2}\, dx$$

$$= \frac{2}{3} \pi \left(\frac{1}{4} x (R^2 - x^2)^{3/2} + \frac{3}{8} R^2 x (R^2 - x^2)^{1/2} + \frac{3}{8} R^4 \operatorname{arcsen} \frac{x}{R} \right)_{-R}^{R}$$

$$= \frac{2}{3} \pi \left(\frac{3}{8} R^4 \pi \right) = \frac{1}{4} \pi^2 R^4$$ ■

Ejemplo 3. Calculemos la integral triple de la función $f(x, y, z) = 1$ sobre la región Ω limitada por los paraboloides $z = x^2 + y^2$, $z = 4x^2 + 4y^2$, el cilindro $y = x^2$ y el plano $y = 3x$. Obsérvese que esta región está limitada por debajo por el paraboloide $z = x^2 + y^2$, por arriba por el paraboloide $z = 4x^2 + 4y^2$, y que su proyección sobre el plano xy está limitada por la parábola $y = x^2$ y la recta $y = 3x$ (las cuales se cortan en los puntos $(0, 0)$ y $(3, 9)$). Así pues, la proyección de Ω sobre el plano xy es la región R dada por

$$R = \{(x, y) | 0 \leq x \leq 3, x^2 \leq y \leq 3x\}$$

de modo que la integral triple que queremos calcular se ve como

$$\iiint_{\Omega} dx\,dy\,dz = \int_0^3 \left(\int_{x^2}^{3x} \left[\int_{x^2+y^2}^{4x^2+4y^2} dz \right] dy \right) dx$$
$$= \int_0^3 \left(\int_{x^2}^{3x} \left[3x^2 + 3y^2 \right] dy \right) dx$$
$$= \int_0^3 3x^2 \left(\int_{x^2}^{3x} dy \right) dx + \int_0^3 \left(\int_{x^2}^{3x} 3y^2\, dy \right) dx$$
$$= \int_0^3 3x^2(3x - x^2)\, dx + \int_0^3 (9x^3 - x^6)\, dx = \frac{9477}{35}$$

■

Ejercicios (Capítulo 6, Sección 6)

1. Para calcular la integral triple de la función continua $f(x, y, z)$ sobre la región $R = \{(x, y, z) | x^2 + y^2 + z^2 \leq 1\}$, se puede proceder a integrar:

a. primero respecto de x, luego respecto de y, por último respecto de z.

b. primero respecto de x, luego respecto de z, por último respecto de y.

c. primero respecto de y, luego respecto de x, por último respecto de z.

d. primero respecto de y, luego respecto de z, por último respecto de x.

e. primero respecto de z, luego respecto de x, por último respecto de y.

f. primero respecto de z, luego respecto de y, por último respecto de x.

En cada caso, escriba la expresión de la integral iterada correspondiente.

2. Repita el ejercicio anterior con la región R limitada por los planos coordenados y el plano $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, $(a, b, c > 0)$.

3. ¿Qué posibilidades hay en el orden de integración para calcular la integral triple de una función $f(x, y, z)$ sobre el cilindro

$$\Omega = \{(x, y, z) | x^2 + y^2 \leq 1, -1 \leq z \leq 1\}?$$

4. ¿Qué posibilidades hay en el orden de integración para calcular la integral triple de una función $f(x, y, z)$ sobre la región

$$\Omega = \{(x, y, z) | x^2 + z^2 \leq 4, -2 \leq y \leq 2\}?$$

5. Escriba la expresión de la integral triple de la función $f(x, y, z)$ sobre el cubo $|x| + |y| + |z| \leq 1$, en términos de integrales iteradas.

En los ejercicios 6—13, calcular la integral triple indicada.

6. $\iiint_{\Omega} xyz\, dx\, dy\, dz$, donde $\Omega = [0, 1] \times [0, 1] \times [0, 1]$.

7. $\iiint_{\Omega} (x + y + z)\, dx\, dy\, dz$, donde $\Omega = [0, 1] \times [0, 1] \times [0, 1]$.

8. $\iiint_{\Omega} (1 - x - y - z)\, dx\, dy\, dz$, donde Ω es la región limitada por los planos coordenados y el plano $x + y + z = 1$.

9. $\iiint_{\Omega} z(x - 1)(y - 2)\, dx\, dy\, dz$, donde Ω es la región del ejercicio anterior.

10. $\iiint_{\Omega} \frac{1}{(1 + x + y + z)^3}\, dx\, dy\, dz$, donde Ω es la parte de la región del ejercicio 5 que se encuentra en el primer octante.

11. $\iiint_{\Omega} e^{x+y+z}\, dx\, dy\, dz$, donde $\Omega = [-1, 1] \times [-1, 1] \times [-1, 1]$.

12. $\iiint_{\Omega} (xe^y + ye^z)\, dx\, dy\, dz$, donde Ω es la región limitada por los planos $x = 0$, $y = 0$, $z = 0$, $y = 1 - x$, $z = 1$.

13. $\iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz$, donde

$$f(x, y, z) = \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2}}$$

y Ω es la región limitada por el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. (Nota: este ejercicio se resolverá fácilmente con la técnica que se estudiará en la siguiente sección).

6.7 Cambio de variables en integrales triples

Como sucede con las integrales dobles, la técnica de introducir nuevas variables en una integral triple resulta ser un instrumento muy poderoso que permite hacer el cálculo de la integral de una manera muy sencilla. En esta sección presentamos el resultado correspondiente al cambio de variables en integrales triples, el cual, como era de esperarse, es completamente análogo al estudiado en la sección 4 para integrales dobles.

La idea general será cambiar las variables x, y, z de la función f en la integral triple

$$\iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz$$

la cual se hace sobre la región Ω de $\mathbb{R}^3$ (donde los puntos se localizan con las coordenadas (x, y, z); diremos que Ω es una región del "espacio xyz") por otras tres nuevas variables, que llamaremos u, v, w, de modo que se calcule la integral anterior integrando una nueva función $\Phi(u, v, w)$ ahora sobre una nueva región Ω' (del espacio $\mathbb{R}^3$ en donde los puntos se localizan con sus coordenadas (u, v, w), que llamaremos "espacio uvw").

Consideremos entonces una función $F: \Omega' \subset \mathbb{R}^3 \to \mathbb{R}^3$ del tipo

$$F(u, v, w) = (x, y, z) = (x(u, v, w), y(u, v, w), z(u, v, w))$$

(llamada función de transformación de coordenadas) que manda a los puntos (u, v, w) de la región Ω' del espacio uvw en puntos $F(u, v, w) = (x, y, z)$ de una región $F(\Omega') = \Omega$ del espacio xyz. Como en el caso de las integrales dobles, se pide que esta función sea de clase $\mathscr{C}^1$, inyectiva y que su jacobiano no se anule en Ω'. En este caso el jacobiano de esta función es

$$\det F'(u, v, w) = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix}$$

Con tales condiciones se tiene la fórmula de cambio de variables en integrales triples

$$\iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iiint_{\Omega'} f(F(u, v, w)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du\, dv\, dw$$

Hay dos casos particulares de cambio de variables en las integrales triples, que estudiaremos con detalle, en los que se trata de cambiar nuestro sistema coordenado cartesiano xyz por el sistema de coordenadas cilíndricas o el de coordenadas esféricas. Antes de estudiar estos sistemas, veamos un par de ejemplos de aplicación de la fórmula anterior.

Ejemplo 1. Se quiere calcular la integral triple de la función $f(x, y, z) = xyz$ sobre la región Ω que se encuentra en el primer octante ($x > 0, y > 0, z > 0$) limitada por los paraboloides $z = x^2 + y^2$, $z = 2x^2 + 2y^2$, por los cilindros $xy = 1$, $xy = 4$, y por los planos $y = x$, $y = 5x$. Es claro que la proyección de esta región en el plano xy es la región del primer cuadrante limitada por las dos hipérbolas $xy = 1$, $xy = 4$, y por las rectas $y = x$, $y = 5x$ (esta es una región difícil de describir, pues es la unión de 3 subregiones del tipo I —o del tipo II). Por otra parte es claro también que Ω está limitada por debajo por el paraboloide $z = x^2 + y^2$ y por arriba por el paraboloide $z = 2x^2 + 2y^2$. Dejamos al lector que verifique que el cálculo de la integral triple en coordenadas cartesianas se vería como

$$\int_{1/\sqrt{5}}^{2/\sqrt{5}} \left(\int_{1/x}^{5x} \left[\int_{x^2+y^2}^{2(x^2+y^2)} xyz\, dz \right] dy \right) dx + \int_{2/\sqrt{5}}^{1} \left(\int_{1/x}^{4/x} \left[\int_{x^2+y^2}^{2x^2+y^2} xyz\, dz \right] dy \right) dx$$
$$+ \int_{1}^{2} \left(\int_{x}^{4/x} \left[\int_{x^2+y^2}^{2(x^2+y^2)} xyz\, dz \right] dy \right) dx$$

Si introducimos las nuevas variables u, v, w tales que

$$u = \frac{z}{x^2 + y^2}, \quad v = xy, \quad w = \frac{y}{x}$$

la región Ω' del espacio uvw que corresponde a nuestra región Ω se describe de una manera muy sencilla. De hecho

$$\Omega' = \{(u, v, w) | 1 \leq u \leq 2, 1 \leq v \leq 4, 1 \leq w \leq 5\}$$

Podemos hacer explícita la función de transformación de coordenadas $F(u, v, w) = (x, y, z)$, resolviendo para x, y, z en términos de u, v, w las expresiones previamente establecidas de las funciones $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$ (de otro modo: se han establecido las funciones coordenadas de F^{-1} y queremos dejar establecidas las funciones coordenadas de F) Puesto que estamos en el primer octante del espacio xyz, se obtiene que

$$x = \sqrt{v/w}, \quad y = \sqrt{vw}, \quad z = \frac{uv(1 + w^2)}{w}$$

Es decir que $F(u, v, w) = \left(\sqrt{v/w}, \sqrt{vw}, \frac{uv(1+w^2)}{w}\right)$. El jacobiano de esta transformación es

$$\begin{aligned}
\frac{\partial(x, y, z)}{\partial(u, v, w)} &= \det \begin{bmatrix} 0 & 0.5(vw)^{-1/2} & -0.5v^{1/2}w^{-3/2} \\ 0 & 0.5(w/v)^{1/2} & 0.5(v/w)^{1/2} \\ \frac{v(1 + w^2)}{w} & \frac{u(1 + w^2)}{w} & \frac{uv(1 - w^2)}{w^2} \end{bmatrix} \\
&= \frac{v}{2w^2}(1 + w^2)
\end{aligned}$$

Así pues, haciendo el cambio de coordenadas en la integral triple nos queda

$$\begin{aligned}
\iiint_{\Omega} xyz\, dx\, dy\, dz &= \iiint_{\Omega'} \sqrt{v/w}\sqrt{vw}\frac{uv(1 + w^2)}{w} \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du\, dv\, dw \\
&= \int_1^2 \left(\int_1^4 \left[\int_1^5 \sqrt{v/w}\sqrt{vw}\frac{uv(1 + w^2)}{w} \frac{v}{2w^2}(1 + w^2)\, dw \right] dv \right) du \\
&= \int_1^2 \left(\int_1^4 \left[\int_1^5 \frac{uv^3(1 + w^2)}{2w^3}\, dw \right] dv \right) du \\
&= \frac{1}{2} \int_1^2 u\, du \int_1^4 v^3\, dv \int_1^5 (w^{-3} + 2w^{-1} + w)\, dw \\
&= \frac{765}{8}\left(\ln 5 + \frac{156}{25} \right)
\end{aligned}$$

■

Ejemplo 2. Consideremos los 6 planos en $\mathbb{R}^3$

$$a_1x + b_1y + c_1z = \alpha_1, \quad a_1x + b_1y + c_1z = \alpha_2 \qquad (\Pi_1)$$

$$a_2x + b_2y + c_2z = \beta_1, \quad a_2x + b_2y + c_2z = \beta_2 \qquad (\Pi_2)$$

$$a_3x + b_3y + c_3z = \gamma_1, \quad a_3x + b_3y + c_3z = \gamma_2 \qquad (\Pi_3)$$

donde $\alpha_1 < \alpha_2, \beta_1 < \beta_2, \gamma_1 < \gamma_2$, y

$$\Delta = \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \neq 0$$

Esta última condición asegura que los planos Π_i *no son paralelos a los planos* Π_j *para* $i \neq j$ (de modo equivalente: los vectores (a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3) que son normales a cada pareja de planos Π_1, Π_2, Π_3, respectivamente, son linealmente independientes —ver la sección 6 del capítulo 1). Se tienen entonces 3 parejas de planos, donde cada pareja consta de planos paralelos, y los planos de cada pareja son no paralelos a los planos de las restantes dos parejas. Estos planos forman entonces un paralelepípedo, que limita una región Ω en $\mathbb{R}^3$. Hay que, integrar la función $f(x, y, z) = 1$ sobre la región Ω (en la próxima sección veremos que lo que se calculará es el volumen del paralelepípedo). Este problema se puede resolver directamente en las coordenadas cartesianas, pero los cálculos son muy complicados (sugerimos al lector que bosqueje el planteamiento del problema de esta manera). Sin embargo, si introducimos las nuevas variables u, v, w de modo que

$$u = a_1x + b_1y + c_1z, \quad v = a_2x + b_2y + c_2z, \quad w = a_3x + b_3y + c_3z$$

la región de integración Ω' en el espacio uvw que corresponde a Ω sería el paralelepípedo

$$\Omega' = \{(u, v, w) | \alpha_1 < u < \alpha_2, \beta_1 < v < \beta_2, \gamma_1 < w < \gamma_2\}$$

Para calcular el jacobiano de la función de transformación de coordenadas $F: \Omega' \to \mathbb{R}^3$, $F(u, v, w) = (x, y, z)$, podemos hacerlo tomando el inverso (algebraico) del jacobiano de la transformación inversa $F^{-1}: \Omega \to \mathbb{R}^3$, $F^{-1}(x, y, z) = (u, v, w) = (a_1x + b_1y + c_1z, a_2x + b_2y + c_2z, a_3x + b_3y + c_3z)$ el cual es

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \det \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix} = \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \Delta$$

(que suponemos distinto de cero). Entonces

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \frac{1}{\frac{\partial(u, v, w)}{\partial(x, y, z)}} = \frac{1}{\Delta}$$

de modo que

$$\begin{aligned} \iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz &= \iiint_{\Omega} dx\, dy\, dz \\ &= \iiint_{\Omega'} \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du\, dv\, dw = \iiint_{\Omega'} \frac{1}{\Delta}\, du\, dv\, dw \\ &= \frac{1}{\Delta} \int_{\alpha_1}^{\alpha_2} \left(\int_{\beta_1}^{\beta_2} \left[\int_{\gamma_1}^{\gamma_2} dw \right] dv \right) du = \frac{1}{\Delta}(\alpha_2 - \alpha_1)(\beta_2 - \beta_1)(\gamma_2 - \gamma_1) \end{aligned}$$ ■

6.7.1 Coordenadas cilíndricas

Introducimos tres nuevas coordenadas para un punto $P = (x, y, z)$ en $\mathbb{R}^3$, denotadas por r, θ, z, según las fórmulas

$$x = r\cos\theta, \quad y = r\operatorname{sen}\theta, \quad z = z$$

(ver apéndice de la sección 5 del capítulo 1) o, en forma equivalente, consideramos la función de transformación de coordenadas $F: \Omega' \to \mathbb{R}^3$, dada por

$$F(r, \theta, z) = (x, y, z) = (r\cos\theta, r\operatorname{sen}\theta, z)$$

A la terna ordenada (r, θ, z) se le llama *coordenadas cilíndricas del punto P*. Nótese que la tercera coordenada z del sistema cartesiano es *la misma* que la tercera coordenada del sistema de coordenadas cilíndricas (que denotamos con la misma letra z). Para que la función F sea inyectiva, se debe pedir que su dominio sea $[0, +\infty) \times [0, 2\pi) \times \mathbb{R}$. Es decir, el espacio $\mathbb{R}^3$ queda cubierto con las coordenadas cilídricas (r, θ, z) con

$$0 \leq r < +\infty, \quad 0 \leq \theta < 2\pi, \quad -\infty < z < \infty$$

Nótese además que las dos primeras coordenadas de un punto $P' = (r, \theta, z)$ en el sistema de coordenadas cilídricas, corresponden a las coordenadas polares de la proyección en el plano xy del correspondiente punto $P = (x, y, z)$ en el sistema cartesiano, en tanto que la tercera coordenada queda igual. En otras palabras, para cambiar del punto $P = (x, y, z)$ en el sistema cartesiano al sistema cilíndrico, consideramos las coordenadas polares del punto $\tilde{P} = (x, y, 0) =$ proyección de P al plano xy, quedando entonces que $x = r\cos\theta$, $y = r\operatorname{sen}\theta$ (estas son las fórmulas de transformación del punto (x, y) en el plano al sistema polar), y dejamos la coordenada z sin variar. Esto se ve así en un esquema

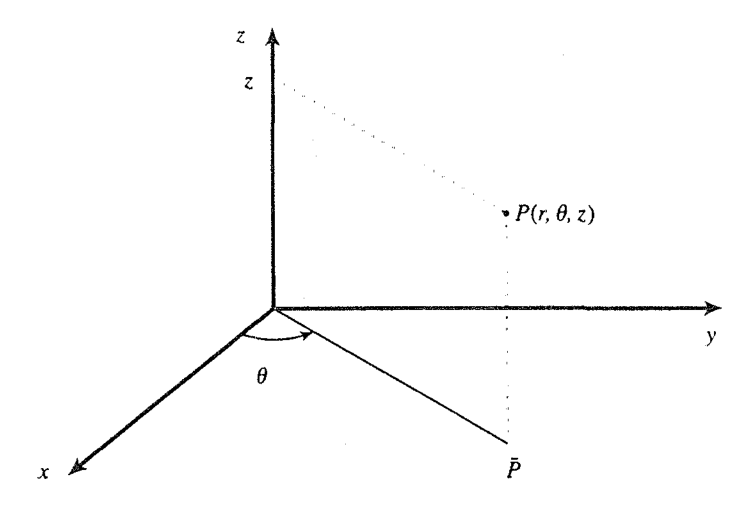


Figura 1. Sistema de coordenadas cilíndricas.

Entonces, la ecuación del cilindro $x^2 + y^2 = R^2$ (que es un cilindro circular recto con el eje z como eje de simetría), se ve, en el sistema cilíndrico, como $r = R$. Por otra parte, en este sistema las ecuaciones del tipo $\theta = cte.$ corresponden a planos perpendiculares al plano xy que contienen al

eje z, es decir, planos de la forma $Ax + By = 0$ (¿por qué?). Por último, es claro que los planos del tipo $z = cte.$ se ven igual en los dos sistemas coordenados (planos paralelos al plano xy, o al $r\theta$ en su caso). De modo pues que los paralelepípedos rectangulares en el espacio $r\theta z$ del tipo

$$\Omega' = \{(r, \theta, z) | R_1 \leq r \leq R_2, \theta_1 \leq \theta \leq \theta_2, z_1 \leq z \leq z_2\}$$

se transforman inyectivamente por la función de transformación de coordenadas $F(r, \theta, z) = (r\cos\theta, r\operatorname{sen}\theta, z)$, en paralelepípedos "cilíndricos", como se muestra en la figura 2.

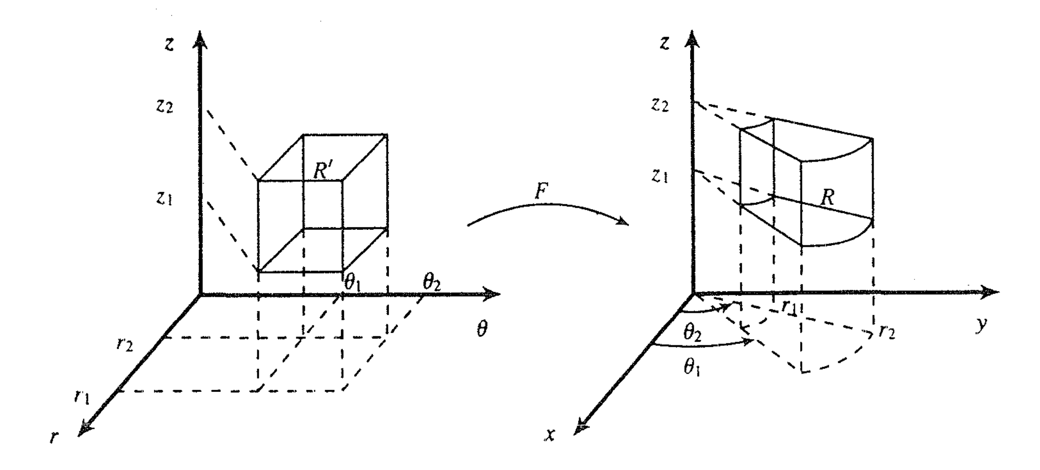


Figura 2. La imagen de un paralelepípedo rectangular bajo la función F.

El jacobiano de la transformación $F(r, \theta, z) = (r\cos\theta, r\operatorname{sen}\theta, z)$ es, en este caso,

$$\det F(r, \theta, z) = \frac{\partial(x, y, z)}{\partial(r, \theta, z)} = \det \begin{bmatrix} \cos\theta & -r\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = r$$

de modo que la fórmula de cambio de variables x, y, z a coordenadas cilíndricas r, θ, z en una integral triple se ve como

$$\iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz = \iiint_{\Omega'} f(r\cos\theta, r\operatorname{sen}\theta, z) r\, dr d\theta dz$$

donde Ω' es la región del espacio $r\theta z$ transformada en Ω por la función F. Aunque no hay una regla general que nos diga cuándo debemos hacer el cambio de variables en una integral triple al sistema cilíndrico, normalmente éste resulta útil cuando la región de integración consta de cilindros y/o planos de los anteriormente mencionados, cuya descripción en coordenadas cilíndricas es muy sencilla.

Ejemplo 3. Calculemos la integral triple de la función $f(x, y, z) = 1 + (x^2 + y^2)^2$ sobre la región Ω limitada por el cono $z = \sqrt{x^2 + y^2}$ y el plano $z = 2$. Aunque este problema se puede

hacer directamente en coordenadas cartesianas, en cuyo caso las operaciones que habría que hacer se verían

$$\int_{-1}^{2}\left(\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\left[\int_{\sqrt{x^2+y^2}}^{2}[1+(x^2+y^2)^2]dz\right]dy\right)dx$$

(verifique e "intente" comenzar éstas) resulta mucho más sencillo si hacemos la transformación a coordenadas cilíndricas. En tal caso, la función a integrar se ve como $f(r\cos\theta, r\operatorname{sen}\theta, z) = 1+(r^2)^2 = 1+r^4$. La región de integración queda limitada en la coordenada z por el cono $z=\sqrt{x^2+y^2}$ cuya ecuación se ve como $z=r$, y por el plano $z=2$ (cuya ecuación se ve igual); la proyección de la región Ω en el plano xy, que es $x^2+y^2\leq 4$ (interior de la intersección del cono con el plano), se describe como $0\leq r\leq 2$, $0\leq\theta\leq 2\pi$. Entonces la integral procurada es

$$\begin{aligned}\int_0^{2\pi}\left(\int_0^2\left[\int_r^2(1+r^4)r\,dz\right]dr\right)d\theta &= \int_0^{2\pi}\left(\int_0^2(1+r^4)r(2-r)dr\right)d\theta\\ &= \int_0^{2\pi}\left(\int_0^2(2r-r^2+2r^5-r^6)dr\right)d\theta = \frac{92}{21}\pi\end{aligned}$$

■

Ejemplo 4. Retomemos el ejemplo 3 de la sección anterior, donde se calculó la integral triple

$$\int_0^3\left(\int_{x^2}^{3x}\left[\int_{x^2+y^2}^{4x^2+4y^2}dz\right]dy\right)dx$$

que corresponde a la integral de la función $f(x,y,z)=1$ sobre la región Ω limitada por los paraboloides $z=x^2+y^2$, $z=4x^2+4y^2$, el cilindro $y=x^2$ y el plano $y=3x$. Hagamos el cálculo de esta integral en coordenadas cilíndricas. Los paraboloides que limitan la región por debajo y por arriba se ven en coordenadas cilíndricas como $z=r^2$ y $z=4r^2$, respectivamente. La proyección de la región de integración sobre el plano xy es la región comprendida entre la parábola $y=x^2$, que en coordenadas cilíndricas se ve como $r\operatorname{sen}\theta = r^2\cos^2\theta$, o sea $r=\frac{\operatorname{sen}\theta}{\cos^2\theta}$, y la recta $y=3x$, que en las nuevas coordenadas es $r\operatorname{sen}\theta = 3r\cos\theta$, o sea $\theta=\arctan 3$. De modo pues, entonces que la región de integración Ω se describe en coordenadas cilíndricas como

$$\Omega' = \{(r,\theta,z)|0\leq r\leq\frac{\operatorname{sen}\theta}{\cos^2\theta}, 0\leq\theta\leq\arctan 3, r^2\leq z\leq 4r^2\}$$

y la integral se calcula como

$$\begin{aligned}\int_0^{\arctan 3}\left(\int_0^{\operatorname{sen}\theta/\cos^2\theta}\left[\int_{r^2}^{4r^2}r\,dz\right]dr\right)d\theta &= \int_0^{\arctan 3}\left(\int_0^{\operatorname{sen}\theta/\cos^2\theta}3r^3\,dr\right)d\theta\\ &= \frac{3}{4}\int_0^{\arctan 3}\frac{\operatorname{sen}^4\theta}{\cos^8\theta}d\theta\\ &= \frac{3}{4}\int_0^{\arctan 3}(\tan^4\theta)(1+\tan^2\theta)\sec^2\theta\,d\theta\\ &= \frac{3}{4}\int_0^3 t^4(1+t^2)\,dt = \frac{9477}{35}\end{aligned}$$

como resultó en el ejemplo 3 de la sección anterior. ■

Una variante de las coordenadas cilíndricas que permite abordar algunos problemas más generales, es introducir, las variables r, θ, $\bar{z}$ en lugar de las variables x, y, z según las fórmulas

$$x = ar\cos\theta, \quad y = br\operatorname{sen}\theta, \quad z = c\bar{z}$$

donde a, b y c son números reales que se escogen de modo conveniente. Se dice que r, θ, $\bar{z}$ son las *coordenadas cilíndricas generalizadas* del punto $P = (x, y, z)$. En estas coordenadas, las expresiones del tipo $\frac{x^2}{a^2} + \frac{y^2}{b^2}$ se ven como r^2, de modo que las ecuaciones de los cilindros elípticos rectos del tipo $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ se ven como $r = 1$.

En este caso el jacobiano de la transformación $F(r, \theta, \bar{z}) = (x, y, z)$ es, como se puede comprobar directamente, el siguiente

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \bar{z})} = \det\begin{bmatrix} a\cos\theta & -ar\operatorname{sen}\theta & 0 \\ b\operatorname{sen}\theta & br\cos\theta & 0 \\ 0 & 0 & c \end{bmatrix} = abcr$$

y la fórmula de cambio de variables a coordenadas cilídricas generalizadas se ve entonces como

$$\iiint_{\Omega} f(x, y, z)\,dx\,dy\,dz = \iiint_{\Omega'} f(ar\cos\theta, br\operatorname{sen}\theta, c\bar{z})abcr\,dr\,d\theta\,d\bar{z}$$

Ejemplo 5. Calculemos la integral triple de la función $f(x, y, z) = k\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right) + 1$ sobre la región Ω de $\mathbb{R}^3$ encerrada por la superficie

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^4}{c^4} = 1$$

Un análisis simple de las simetrías de esta superficie nos muestra que se trata de una superficie cerrada simétrica a todos los ejes y a todos los planos coordenados (la ecuación no sufre modificación alguna al sustituír en ella x por $-x$, y/o y por $-y$, y/o z por $-z$). Pasando las coordenadas cilíndricas generalizadas r, θ, $\bar{z}$ según las fórmulas $x = ar\cos\theta$, $y = br\operatorname{sen}\theta$, $z = c\bar{z}$, la región Ω corresponde a la región Ω' (en el espacio $r\theta\bar{z}$) limitada por la superficie que tiene por ecuación

$$r^2 + \bar{z}^4 = 1, \quad \text{o sea}, \bar{z} = \pm\sqrt[4]{1 - r^2}$$

limitada entonces por $0 \leq r \leq 1$ (obsérvese el dominio de la función $\bar{z} = \bar{z}(r)$), $0 \leq \theta \leq 2\pi$, $-\sqrt[4]{1 - r^2} \leq \bar{z} \leq \sqrt[4]{1 - r^2}$. El nuevo integrando es

$$f(ar\cos\theta, br\operatorname{sen}\theta, c\bar{z}) = kr^2 + 1$$

de modo que la integral requerida es

$$\int_0^{2\pi}\left(\int_0^1\left[\int_{-\sqrt[4]{1-r^2}}^{\sqrt[4]{1-r^2}}(kr^2 + 1)abcr\,d\bar{z}\right]dr\right)d\theta$$

$$= 2\pi abc\int_0^1 2\sqrt[4]{1 - r^2}(kr^2 + 1)r\,dr$$

$$= 4\pi abc\left[k\int_0^1 \sqrt[4]{1-r^2}\,r^3\,dr + \int_0^1 \sqrt[4]{1-r^2}\,r\,dr\right]$$
$$= 4\pi abc\left[\frac{2k}{9}(1-r^2)^{9/4} - \frac{2k}{5}(1-r^2)^{5/4} - \frac{2}{5}(1-r^2)^{5/4}\right]_0^1$$
$$= 4\pi abc\left(\frac{8k+18}{45}\right)$$ ■

6.7.2 Coordenadas esféricas

Introducimos tres nuevas coordenadas para un punto $P = (x, y, z)$ en $\mathbb{R}^3$, denotadas por r, θ, ϕ, según las fórmulas

$$x = r\cos\theta\operatorname{sen}\phi, \quad y = r\operatorname{sen}\theta\operatorname{sen}\phi, \quad z = r\cos\phi$$

(ver apéndice de la sección 5 del capítulo 1) o, de modo equivalente, consideramos la función de transformación de coordenadas $F\colon \Omega' \to \mathbb{R}^3$, dada por

$$F(r, \theta, \phi) = (x, y, z) = (r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$$

A la terna ordenada (r, θ, ϕ) se le llama *coordenadas esféricas del punto P*. Es fácil ver que los parámetros r, θ, ϕ de las fórmulas anteriores corresponden, respectivamente, a la distancia r del punto $P = (x, y, z)$ al origen de coordenadas, al ángulo θ que forma (en el plano xy) el segmento que une el origen con el punto $P' = (x, y, 0)$ = proyección de P en el plano xy, con la parte positiva del eje x, y, al ángulo que forma el segmento que une el origen con el punto P con la parte positiva del eje z, como se muestra en la figura siguiente.

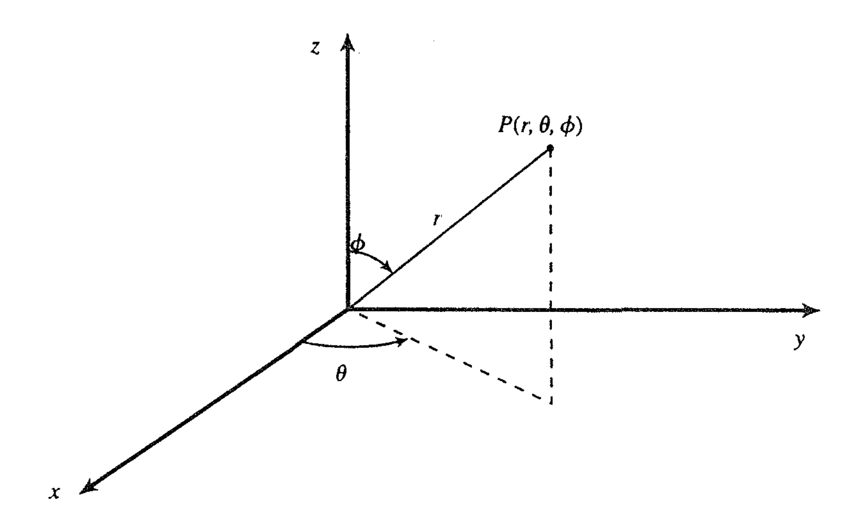


Figura 3. El sistema coordenado esférico.

Para que la función F sea inyectiva, los rangos de variación de las coordenadas esféricas r, θ, ϕ se toman como

$$r \geq 0, \quad 0 \leq \theta < 2\pi, \quad 0 \leq \phi \leq \pi$$

En el sistema de coordenadas esféricas la ecuación de la esfera con centro en el origen y radio R, $x^2 + y^2 + z^2 = R^2$, se ve como

$$\begin{aligned} R^2 &= x^2 + y^2 + z^2 = (r\cos\theta \operatorname{sen}\phi)^2 + (r \operatorname{sen}\theta \operatorname{sen}\phi)^2 + (r\cos\phi)^2 \\ &= r^2 \operatorname{sen}^2\phi(\cos^2\theta + \operatorname{sen}^2\theta) + r^2\cos^2\phi = r^2(\operatorname{sen}^2\phi + \cos^2\phi) = r^2 \end{aligned}$$

es decir, como $r = R$. Por otra parte, puesto que la coordenada θ en el sistema esférico es la misma que en sistema cilíndrico, los planos de la forma $Ax + By = 0$ corresponden a ecuaciones del tipo $\theta = cte$. Por último, las ecuaciones del tipo $\phi = cte$. corresponden a ecuaciones de semiconos (para $\phi \neq \pi/2$; la ecuación $\phi = \pi/2$ es la ecuación del plano xy) del tipo $z = \pm\sqrt{x^2 + y^2}$ (verifique). De tal modo entonces que el paralelepípedo rectangular en el espacio $r\theta\phi$

$$\Omega' = \{(r, \theta, \phi) | R_1 \leq r \leq R_2, \ \theta_1 \leq \theta \leq \theta_2, \ \phi_1 \leq \phi \leq \phi_2\}$$

es transformado, por medio de la función de transformación de coordenadas $F(r, \theta, \phi) = (r\cos\theta \operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$, en un paralelepípedo "esférico" como se muestra en la siguiente figura

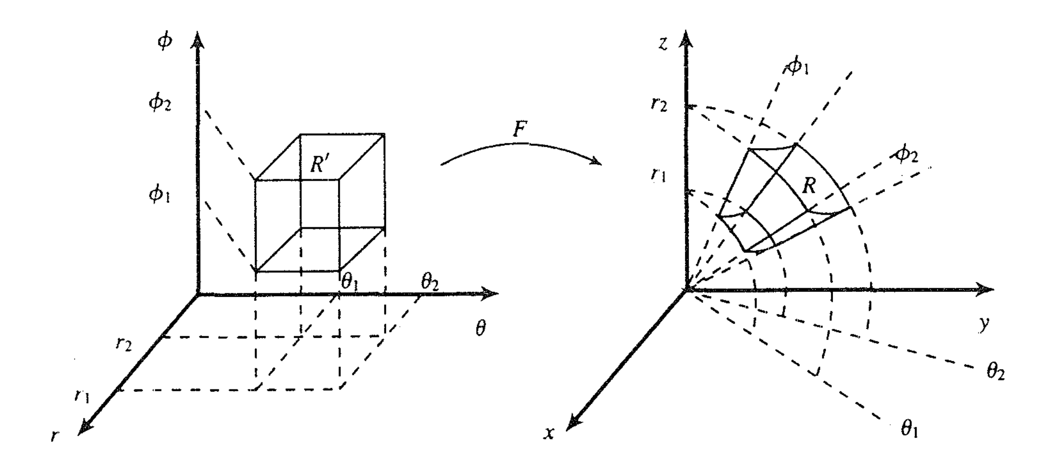


Figura 4. Imagen de un paralelepípedo rectangular bajo la función F.

El jacobiano de la función de transformación a coordenadas esféricas $F(r, \theta, \phi) = (r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$ es

$$\begin{aligned} \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} &= \det\begin{bmatrix} \cos\theta\operatorname{sen}\phi & -r\operatorname{sen}\theta\operatorname{sen}\phi & r\cos\theta\cos\phi \\ \operatorname{sen}\theta\operatorname{sen}\phi & r\cos\theta\operatorname{sen}\phi & r\operatorname{sen}\theta\cos\phi \\ \cos\phi & 0 & -r\operatorname{sen}\phi \end{bmatrix} \\ &= \cos\phi\det\begin{bmatrix} -r\operatorname{sen}\theta\operatorname{sen}\phi & r\cos\theta\cos\phi \\ r\cos\theta\operatorname{sen}\phi & r\operatorname{sen}\theta\cos\phi \end{bmatrix} - r\operatorname{sen}\phi\det\begin{bmatrix} \cos\theta\operatorname{sen}\phi & -r\operatorname{sen}\theta\operatorname{sen}\phi \\ \operatorname{sen}\theta\operatorname{sen}\phi & r\cos\theta\operatorname{sen}\phi \end{bmatrix} \\ &= \cos\phi(-r^2\operatorname{sen}\phi\cos\phi) - r\operatorname{sen}\phi(r\operatorname{sen}^2\phi) = -r^2\operatorname{sen}\phi \end{aligned}$$

de modo que la fórmula de cambio de variables en el caso de las coordenadas esféricas se ve como

$$\iiint_{\Omega} f(x,y,z)\,dx\,dy\,dz = \iiint_{\Omega'} f(r\cos\theta \operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)r^2 \operatorname{sen}\phi\,dr\,d\theta\,d\phi$$

Como en el caso de cualquier cambio de variables, no hay una regla general que nos diga cuándo debemos hacer el cambio a coordenadas esféricas. Sin embargo, podemos decir que éste es recomendable cuando en la región de integración Ω aparezcan esferas o trozos de ellas que se describan fácilmente en este tipo de coordenadas, y/o que en el integrando aparezcan expresiones del tipo $x^2 + y^2 + z^2$, las cuales, en estas coordenadas, se ven como r^2.

Ejemplo 6. Retomemos el ejemplo 2 de la sección anterior. Se quiere calcular la integral

$$\iiint_{\Omega} \sqrt{R^2 - x^2 - y^2 - z^2}\,dx\,dy\,dz$$

donde Ω es la región limitada por la esfera $x^2 + y^2 + z^2 = R^2$. Esta región de integración se ve así en coordenadas esféricas

$$\Omega' = \{(r,\theta,\phi)|0 \le r \le R, 0 \le \theta \le 2\pi, 0 \le \phi \le \pi\}$$

El integrando $f(x,y,z) = \sqrt{R^2 - x^2 - y^2 - z^2}$ se ve en las nuevas coordenadas como

$$f(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi) = \sqrt{R^2 - r^2}$$

de modo que la integral por calcular queda

$$\begin{aligned}
\iiint_{\Omega'} \sqrt{R^2 - r^2}r^2 \operatorname{sen}\phi\,dr\,d\theta\,d\phi &= \int_0^R \left(\int_0^{2\pi}\left[\int_0^{\pi} \sqrt{R^2 - r^2}r^2 \operatorname{sen}\phi\,d\phi\right]d\theta\right)dr \\
&= \int_0^R \sqrt{R^2 - r^2}r^2\,dr \int_0^{2\pi} d\theta \int_0^{\pi} \operatorname{sen}\phi\,d\phi \\
&= 4\pi \int_0^R \sqrt{R^2 - r^2}r^2\,dr \\
&= \text{“}\overset{r = R\operatorname{sen}t}{\longrightarrow}\text{”} = 4\pi \int_0^{\pi/2} (R\cos t)(R\operatorname{sen}t)^2(R\cos t)\,dt \\
&= 4\pi R^4 \int_0^{\pi/2} \cos^2 t\operatorname{sen}^2 t\,dt = \pi R^4 \int_0^{\pi/2} \operatorname{sen}^2 2t\,dt \\
&= \frac{1}{2}\pi R^4 \int_0^{\pi/2} (1 - \cos 4t)\,dt = \frac{1}{2}\pi R^4 \left(t - \frac{1}{4}\operatorname{sen} 4t\right)_0^{\pi/2} \\
&= \frac{1}{4}\pi^2 R^4
\end{aligned}$$

como se obtuvo en el ejemplo 2 de la sección anterior. ■

Ejemplo 7. Calculemos la integral de la función $f(x, y, z) = 1$ sobre la región Ω comprendida por la parte de la esfera $x^2+y^2+z^2 = 2$ que queda dentro del cono $z = \sqrt{x^2 + y^2}$. Aunque el problema se puede plantear y resolver (con muchas dificultades técnicas) directamente en coordenadas cartesianas, quedando

$$\int_{-1}^{1} \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \left[\int_{\sqrt{x^2+y^2}}^{\sqrt{2-x^2-y^2}} dz \right] dy \right) dx$$

(verifique) lo resolveremos pasándolo primeramente a las coordenadas esféricas. La esfera involucrada en Ω se ve como $r = \sqrt{2}$ y el cono se ve como $\phi = \pi/4$. Entonces esta región corresponde en el espacio $r\theta\phi$ a la región Ω' dada por

$$\Omega' = \{(r, \theta, \phi) | 0 \leq r \leq \sqrt{2}, 0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi/4\}$$

y así, la integral por calcular se ve como

$$\begin{aligned} \int_0^{2\pi} \left(\int_0^{\pi/4} \left[\int_0^{\sqrt{2}} r^2 \operatorname{sen} \phi \, dr \right] d\phi \right) d\theta &= \int_0^{2\pi} \left(\int_0^{\pi/4} \operatorname{sen} \phi \, d\phi \left[\frac{1}{3} 2^{3/2} \right] \right) \\ &= \frac{2\pi}{3} 2^{3/2} (\cos \phi)_0^{\pi/4} = \frac{\pi}{3} 2^{3/2} (2 - \sqrt{2}) \end{aligned}$$ ■

En algunos problemas puede ser útil introducir las *coordenadas esféricas generalizadas* r, θ, ϕ, según las fórmulas

$$\begin{aligned} x &= ar \cos^\alpha \theta \operatorname{sen}^\beta \phi \\ y &= br \operatorname{sen}^\alpha \theta \operatorname{sen}^\beta \phi \\ z &= cr \cos^\beta \phi \end{aligned}$$

en las que a, b, c, α, β son números reales que se escogen según convenga. En este caso el jacobiano de la función de transformación a estas coordenadas es

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = -abc\alpha\beta r^2 \cos^{\alpha-1} \theta \operatorname{sen}^{\alpha-1} \theta \cos^{\beta-1} \phi \operatorname{sen}^{2\beta-1} \phi$$

como el lector puede comprobar directamente.

Ejemplo 8. Calculemos la integral triple de la función $f(x, y, z) = 1$ sobre la región Ω en el primer octante limitada por la superficie

$$\left(\frac{x}{a} + \frac{y}{b} + \frac{z}{c} \right)^4 = \frac{xyz}{abc}$$

(a, b y c son números positivos dados). Al pasar a las coordenadas esféricas generalizadas, según las fórmulas

$$x = ar \cos^2 \theta \operatorname{sen}^2 \phi, \quad y = br \operatorname{sen}^2 \theta \operatorname{sen}^2 \phi, \quad z = cr \cos^2 \phi$$

logramos que la expresión $\frac{x}{a} + \frac{y}{b} + \frac{z}{c}$ se convierta en r, de modo que la superficie que limita la región de integración se ve como

$$r = \cos^2 \theta \operatorname{sen}^2 \theta \cos^2 \phi \operatorname{sen}^4 \phi$$

queda entonces la nueva región de integración Ω' en el espacio $r\theta\phi$ como

$$\Omega' = \{(r, \theta, \phi)|0 \leq r \leq \cos^2\theta \operatorname{sen}^2\theta \cos^2\phi \operatorname{sen}^4\phi, 0 \leq \theta \leq \pi/2, 0 \leq \phi \leq \pi/2\}$$

En este caso el valor absoluto del jacobiano de la transformación es

$$\left|\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)}\right| = 4abcr^2 \cos\theta \operatorname{sen}\theta \cos\phi \operatorname{sen}^3\phi$$

de modo que la integral triple por calcular queda como

$$\begin{aligned}
&\int_0^{\pi/2} \left(\int_0^{\pi/2} \left[\int_0^{\cos^2\theta \operatorname{sen}^2\theta \cos^2\phi \operatorname{sen}^4\phi} 4abcr^2 \cos\theta \operatorname{sen}\theta \cos\phi \operatorname{sen}^3\phi \, dr \right] d\theta \right) d\phi \\
&= \frac{4}{3}abc \int_0^{\pi/2} \left(\int_0^{\pi/2} (\cos^2\theta \operatorname{sen}^2\theta \cos^2\phi \operatorname{sen}^4\phi)^3 \cos\theta \operatorname{sen}\theta \cos\phi \operatorname{sen}^3\phi \, d\theta \right) d\phi \\
&= \frac{4}{3}abc \int_0^{\pi/2} \left(\int_0^{\pi/2} \cos^7\theta \operatorname{sen}^7\theta \cos^7\phi \operatorname{sen}^{15}\phi \, d\theta \right) d\phi \\
&= \frac{4}{3}abc \int_0^{\pi/2} \cos^7\phi \operatorname{sen}^{15}\phi \, d\phi \int_0^{\pi/2} cos^7\theta \operatorname{sen}^7\theta \, d\theta \\
&= \frac{4}{3}abc \int_0^{\pi/2} \left((1 - \operatorname{sen}^2\phi)^3 \operatorname{sen}^{15}\phi\right) \cos\phi \, d\phi \int_0^{\pi/2} \left((1 - \operatorname{sen}^2\theta)^3 \operatorname{sen}^7\theta\right) \cos\theta \, d\theta \\
&= \frac{4}{3}abc \left(\frac{1}{2640}\right)\left(\frac{1}{280}\right) \\
&= \frac{abc}{554400}
\end{aligned}$$

■

Ejercicios (Capítulo 6, Sección 7)

En los ejercicios 1—5, se dan regiones Ω en el espacio $\mathbb{R}^3$. Describa estas regiones en el sistema de coordenadas cilíndricas.

1. $\Omega = \{(x, y, z)|x^2 + y^2 \leq 1, x \geq 0, y \geq 0, -1 \leq z \leq 1\}$

2. $\Omega = \{(x, y, z)|x^2 + y^2 \leq 1, x \geq 0, y \leq 0, -1 \leq z \leq 1\}$

3. $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 \leq 4, 0 \leq z \leq 1\}$

4. $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 \leq 4, x \geq 0, x \leq y \leq \sqrt{3}x, 0 \leq z \leq 2\}$

5. $\Omega = \{(x, y, z)|x^2 + y^2 \leq 2x, 1 \leq z \leq 4\}$

6. $\Omega = \{(x, y, z)|x^2 + y^2 \leq 1, z \leq 2x^2 + 2y^2\}$

7. $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 \leq 9, x^2 + y^2 \leq z \leq 3x^2 + 3y^2\}$

En los ejercicios 8—15, se dan regiones Ω en el espacio $\mathbb{R}^3$. Describa estas regiones en el sistema de coordenadas esféricas.

8. $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 1, z \geq 0\}$

9. $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 1, z \leq 0\}$

10. $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 4, x \leq 0, y \geq 0, z \geq 0\}$

11. $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 + z^2 \leq 4, x \geq 0, y \leq 0, z \leq 0\}$

12. $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 + z^2 \leq 9, z \leq 0\}$

13. $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 4, z^2 \geq x^2 + y^2, z \geq 0\}$

14. $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 4, z^2 \leq x^2 + y^2, z \leq 0\}$

15. $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 + z^2 \leq 9, x \leq y \leq \sqrt{3}x\}$

En los ejercicios 16—25, calcule las integrales triples indicadas.

16. $\iiint_\Omega (x^2 + y^2)\, dx\, dy\, dz$, donde Ω es la región del ejercicio 1.

17. $\iiint_\Omega xyz\, dx\, dy\, dz$, donde Ω es la región del ejercicio 2.

18. $\iiint_\Omega z^2 \operatorname{sen}(x^2 + y^2)\, dx\, dy\, dz$, donde Ω es la región del ejercicio 3.

19. $\iiint_\Omega z\sqrt{1 - x^2 - y^2}\, dx\, dy\, dz$, donde Ω es la región del ejercicio 6.

20. $\iiint_\Omega (x^2 + y^2 + z^2) dx\, dy\, dz$, donde Ω es la región limitada por el cilindro $x^2 + y^2 = 16$, $3 \leq z \leq 4$.

21. $\iiint_\Omega (x^2 + y^2 + z^2) dx\, dy\, dz$, donde Ω es la región limitada por la esfera $x^2 + y^2 + z^2 = 1$.

22. $\iiint_\Omega (x^2 + y^2) dx\, dy\, dz$, donde Ω es la región del ejercicio 8.

23. $\iiint_\Omega \frac{1}{1 + x^2 + y^2 + z^2} dx\, dydz$, donde Ω es la región del ejercicio 9.

24. $\iiint_\Omega (x + 2z) dx\, dy\, dz$, donde Ω es la región del ejercicio 12.

25. $\iiint_\Omega z\, dx\, dy\, dz$, donde Ω es la parte común de las esferas $x^2+y^2+z^2 \leq 1$, $x^2+y^2+(z-1)^2 \leq 1$.

6.8 Aplicaciones de las integrales triples

En esta sección presentamos algunas aplicaciones de las integrales triples en problemas de geometría y de mecánica.

6.8.1 Volúmenes de cuerpos en el espacio

Al calcular la integral triple de una función continua positiva $f(x, y, z)$ sobre una región Ω de $\mathbb{R}^3$ se está calculando un "hipervolumen bajo la gráfica de la función sobre Ω" (recuerde que la gráfica de la función f, es decir el conjunto $\{(x, y, z, u)|(x, y, z) \in$ dominio de f, $u = f(x, y, z)\}$ vive en el espacio $\mathbb{R}^4$). Esto se ve claro si recordamos cómo se calcula la integral triple de una función constante, digamos $f(x, y, z) = k$ (un caso concreto de función escalonada), sobre el rectángulo $Q = [a, b] \times [c, d] \times [e, g]$ en $\mathbb{R}^3$, el cual se divide en $m_1 m_2 m_3$ subrectángulos Q_{ijk}. Esta sería (ver fórmula previa a la definición en la sección 6)

$$\begin{aligned}\iiint_Q f(x, y, z)\, dx\, dy\, dz &= \sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\sum_{k=1}^{m_3}\big(\text{valor de } f \text{ en } Q_{ijk}\big)\big(\text{volumen de } Q_{ijk}\big)\\ &= \sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\sum_{k=1}^{m_3}(k)\big(\text{ volumen de } Q_{ijk}\big) = k\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\sum_{k=1}^{m_3}\big(\text{volumen de } Q_{ijk}\big)\\ &= k(\text{volumen de } Q)\end{aligned}$$

De aquí se puede ver que si $k = 1$, el valor de la integral triple de f sobre Q es justamente el volumen de Q. Este resultado también es cierto para regiones más generales Ω en $\mathbb{R}^3$. De hecho se tiene

$$\iiint_\Omega dx\, dy\, dz = \text{ volumen de } \Omega$$

el cual es un resultado análogo al establecido en la sección 5 con integrales dobles, en el que $\iint_R dx\, dy =$ área de R. Advertimos que la fórmula anterior establece una igualdad *numérica*, pues en realidad el resultado del cálculo de la integral triple de f sobre Ω tiene por unidades a unidades de longitud a la cuarta potencia.

Por ejemplo, si la región Ω está limitada debajo por la gráfica de la función $z = f_1(x, y)$ y arriba por la gráfica de la función $z = f_2(x, y)$, y su proyección sobre el plano xy es la región R, es decir si Ω es

$$\Omega = \{(x, y, z)|(x, y) \in R,\ f_1(x, y) \leq z \leq f_2(x, y)\}$$

entonces el volumen de Ω sería

$$\begin{aligned}\iiint_\Omega dx\, dy\, dz &= \iint_R \left(\int_{f_1(x,y)}^{f_2(x,y)} dz\right) dx\, dy\\ &= \iint_R \big(f_2(x, y) - f_1(x, y)\big) dx\, dy\end{aligned}$$

lo cual coincide con la manera como calculamos volúmenes de cuerpos en $\mathbb{R}^3$ con integrales dobles en la sección 5. Por supuesto que la parte interesante en el cálculo de volúmenes con integrales triples, es cuando éstas se calculan con los cambios de variables estudiados en la sección anterior.

Ejemplo 1. En el ejemplo 3 de la sección 6 se calculó la integral triple de la función $f(x, y, z) = 1$ sobre la región Ω limitada por los paraboloides $z = x^2 + y^2$, $z = 4x^2 + 4y^2$, el cilindro $y = x^2$ y el plano $y = 3x$ (ver también el ejemplo 4 de la sección 7). Se obtuvo que

$$\iiint_{\Omega} dx\,dy\,dz = \frac{9477}{35}$$

Este es entonces el volumen de la región Ω. ∎

Ejemplo 2. En el ejemplo 2 de la sección 7 se calculó la integral

$$\iiint_{\Omega} dx\,dy\,dz$$

sobre el paralelepípedo Ω limitado por los 6 planos en $\mathbb{R}^3$

$$a_1x + b_1y + c_1z = \alpha_1, \quad a_1x + b_1y + c_1z = \alpha_2 \qquad (\Pi_1)$$
$$a_2x + b_2y + c_2z = \beta_1, \quad a_2x + b_2y + c_2z = \beta_2 \qquad (\Pi_2)$$
$$a_3x + b_3y + c_3z = \gamma_1, \quad a_3x + b_3y + c_3z = \gamma_2 \qquad (\Pi_3)$$

donde $\alpha_1 < \alpha_2$, $\beta_1 < \beta_2$, $\gamma_1 < \gamma_2$, y

$$\Delta = \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \neq 0$$

Se obtuvo que

$$\iiint_{\Omega} dx\,dy\,dz = \frac{1}{\Delta}(\alpha_2 - \alpha_1)(\beta_2 - \beta_1)(\gamma_2 - \gamma_1)$$

Este es entonces el volumen del paralelepípedo Ω. ∎

Ejemplo 3. Calculemos el volumen encerrado por el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

(ver el ejemplo 6 de la sección 5). Pasando a las coordenadas esféricas generalizadas, $x = ar\cos\theta\,\mathrm{sen}\,\phi$, $y = br\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\phi$, $z = cr\cos\phi$, la expresión del elipsoide se transforma en $r = 1$, y entonces la región de integración corresponde en estas coordenadas a

$$\Omega' = \{(r, \theta, \phi) | 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi\}$$

por lo que el volumen procurado es

$$\iiint_{\Omega'} abcr^2\,\mathrm{sen}\,\phi\,dr\,d\theta\,d\phi = abc\left(\int_0^{2\pi} d\theta\right)\left(\int_0^{\pi} \mathrm{sen}\,\phi\,d\phi\right)\left(\int_0^1 r^2\,dr\right) = abc(2\pi)(2)(1/3)$$
$$= \frac{4}{3}\pi abc$$ ∎

Ejemplo 4. En el ejemplo 5 de la sección 7 se calculó la integral triple

$$\iiint_{\Omega} \left[k\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)+1\right] dx\, dy\, dz$$

donde Ω era la región en $\mathbb{R}^3$ encerrada por la superficie

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^4}{c^4}=1$$

Se obtuvo que el valor de la integral era

$$4\pi abc\left(\frac{8k+18}{45}\right)$$

Nótese que si hacemos $k=0$, la función en el integrando es $f(x, y, z)=1$, y el resultado nos quedaría como

$$\iiint_{\Omega} dx\, dy\, dz = 4\pi abc\left(\frac{18}{45}\right)=\frac{8}{5}\pi abc$$

Entonces el volumen encerrado por la superficie

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^4}{c^4}=1$$

■

Ejemplo 5. En el ejemplo 8 de la sección 7 se calculó la integral de función $f(x, y, z)=1$ sobre la región Ω en el primer octante limitada por la superficie

$$\left(\frac{x}{a}+\frac{y}{b}+\frac{z}{c}\right)^4=\frac{xyz}{abc}$$

(a, b y c son números positivos dados). El resultado obtenido fue

$$\iiint_{\Omega} dx\, dy\, dz=\frac{abc}{554400}$$

El volumen encerrado por la parte de la superficie

$$\left(\frac{x}{a}+\frac{y}{b}+\frac{z}{c}\right)^4=\frac{xyz}{abc}$$

que se encuentra en el primer octante. ■

Ejemplo 6. Consideremos los dos elipsoides

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{(z-c)^2}{c^2}=1$$

Se quiere calcular el volumen de la parte común de ellos. Restando una ecuación de otra se encuentra que la intersección de los elipsoides se realiza en $z=c/2$. La proyección de esta intersección sobre el plano xy es entonces (sustituyendo $z=c/2$ en cualquiera de las dos ecuaciones)

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{3}{4}$$

Dejamos al lector verificar que en coordenadas cartesianas el cálculo del volumen requerido se ve como

$$\int_{-\sqrt{3}a/2}^{\sqrt{3}a/2}\left(\int_{-b\sqrt{3/4-x^2/a^2}}^{b\sqrt{3/4-x^2/a^2}}\left[\int_{c(1-\sqrt{1-x^2/a^2-y^2/b^2})}^{c\sqrt{1-x^2/a^2-y^2/b^2}} dz\right] dy\right) dx$$

(el cual se antoja complicado —lo es). Este problema, sin embargo, se puede resolver pasando a las coordenadas esféricas generalizadas, haciendo

$$x=ar\cos\theta\,\mathrm{sen}\,\phi, \quad y=br\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\phi, \quad z=cr\cos\phi$$

El elipsoide $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ se ve en este caso como $r=1$, en tanto que el elipsoide $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{(z-c)^2}{c^2}=1$, se ve como $r=2\cos\phi$. Debemos dividir la región Ω en dos subregiones, divididas por el valor de ϕ donde ocurre la intersección de las superficies (en ese momento cambia la r, de un elipsoide al otro). Esta intersección se puede hallar resolviendo simultáneamente las ecuaciones $r=1$ y $r=2\cos\phi$, obteniendo así $\phi=\arccos 0.5=\pi/3$. Así, la región Ω corresponde a la unión de las dos subregiones Ω_1' y Ω_2' en el espacio $r\theta\phi$, a saber

$$\Omega_1'=\{(r,\theta,\phi)|0\le r\le 2\cos\phi, 0\le\theta\le 2\pi, 0\le\phi\le\pi/3\}$$
$$\Omega_2'=\{(r,\theta,\phi)|0\le r\le 1, 0\le\theta\le 2\pi, \pi/3\le\phi\le\pi/2\}$$

de modo que el volumen requerido es (en el integrando sólo aparecerá el valor absoluto del jacobiano de la tranformación, que en este caso es igual a $abcr^2$ sen ϕ)

$$\iiint_{\Omega_1'} abcr^2\,\mathrm{sen}\,\phi\,dr\,d\theta\,d\phi+\iiint_{\Omega_2'} abcr^2\,\mathrm{sen}\,\phi\,dr\,d\theta\,d\phi$$
$$=\int_0^{2\pi}\left(\int_0^{\pi/3}\left[\int_0^1 abcr^2\,\mathrm{sen}\,\phi\,dr\right]d\phi\right)d\theta$$
$$+\int_0^{2\pi}\left(\int_{\pi/3}^{\pi/2}\left[\int_0^{2\cos\phi} abcr^2\,\mathrm{sen}\,\phi\,dr\right]d\phi\right)d\theta$$
$$=\frac{abc}{3}\left[\int_0^{2\pi}\left(\int_0^{\pi/3}\mathrm{sen}\,\phi\,d\phi\right)d\theta+\int_0^{2\pi}\left(\int_{\pi/3}^{\pi/2}(2\cos\phi)^3\,\mathrm{sen}\,\phi\,d\phi\right)d\theta\right]$$
$$=\frac{2abc\pi}{3}\left[\left(-\cos\phi\right)_0^{\pi/3}+\left(-2\cos^4\phi\right)_{\pi/3}^{\pi/2}\right]=\frac{5}{12}\pi abc$$

■

Ejemplo 7. Se quiere calcular el volumen encerrado por la superficie

$$(x^2 + y^2 + z^2)^2 = x^2 + y^2 - z^2$$

Atendiendo a las simetrías que presenta esta superficie, es claro que se trata de una superficie cerrada simétrica respecto de todos los ejes y planos coordenados. Podemos, por ejemplo, calcular el volumen encerrado por la superficie en el primer octante y multiplicarlo por 8. Además, es claro que en este caso *estamos obligados* a hacer un cambio de variables, pues el problema resulta inmanejable en coordenadas cartesianas. El miembro izquierdo de la ecuación que define a la superficie sugiere que el cambio lo debemos hacer a las coordenadas esféricas, poniendo $x = r\cos\theta\,\text{sen}\,\phi$, $y = r\,\text{sen}\,\theta\,\text{sen}\,\phi$, $z = r\cos\phi$, en cuyo caso la ecuación de la superficie se ve como

$$(r^2)^2 = (r\cos\theta\,\text{sen}\,\phi)^2 + (r\,\text{sen}\,\theta\,\text{sen}\,\phi)^2 - (r\cos\phi)^2$$

o bien, puesto que $r \geq 0$,

$$r = \sqrt{\text{sen}^2\phi - \cos^2\phi} = \sqrt{-\cos 2\phi}$$

Nótese que para que esta función $r = r(\phi)$ exista se debe tener $-\cos 2\phi \geq 0$, o bien $\cos 2\phi \leq 0$, de donde $\pi/4 \leq \phi \leq 3\pi/4$. Como sólo calcularemos el volumen en el primer octante, tomamos ϕ que varía entre $\pi/4$ y $\pi/2$. Entonces la región de integración quedará descrita como

$$\Omega' = \{(r, \theta, \phi) | 0 \leq r \leq \sqrt{-\cos 2\phi}, 0 \leq \theta \leq \pi/2, \pi/4 \leq \phi \leq \pi/2\}$$

y así, el volumen V requerido (solamente integramos el jacobiano de la transformación que es $r^2\,\text{sen}\,\phi$) es

$$\begin{aligned} V &= 8\int_0^{\pi/2}\left(\int_{\pi/4}^{\pi/2}\left[\int_0^{\sqrt{-\cos 2\phi}} r^2\,\text{sen}\,\phi\,dr\right]d\phi\right)d\theta \\ &= \frac{4\pi}{3}\int_{\pi/4}^{\pi/2}(-\cos 2\phi)^{3/2}\,\text{sen}\,\phi\,d\phi = \frac{4\pi}{3}\int_{\pi/4}^{\pi/2}(1 - 2\cos^2\phi)^{3/2}(\text{sen}\,\phi)\,d\phi \\ &= \frac{4\pi}{3}\int_0^{\sqrt{2}/2}(1 - 2u^2)^{3/2}\,du = \frac{4\pi}{3\sqrt{2}}\int_0^{\pi/2}\cos^4\xi\,d\xi = \left(\frac{4\pi}{3\sqrt{2}}\right)\left(\frac{3\pi}{16}\right) \\ &= \frac{\pi^2}{4\sqrt{2}} \end{aligned}$$

■

6.8.2 Centros de masa y momentos de cuerpos en el espacio

Se presentan los mismos conceptos de momento estático y centro de masa, estudiados en la sección 5 para cuerpos planos, considerando ahora cuerpos que ocupan una región Ω en el espacio. Si el cuerpo es no homogéneo, con función densidad $\rho\colon \Omega \to \mathbb{R}$ (en este caso para cada punto $(x, y, z) \in \Omega$, $\rho(x, y, z)$ es el valor de la densidad del cuerpo en ese punto, medida en unidades de masa por unidades de volumen, por ejemplo gr/cm^3). La *masa total del cuerpo* Ω es entonces

$$M = \iiint\limits_{\Omega} \rho(x, y, z)\,dx\,dy\,dz$$

Los momentos estáticos del cuerpo Ω respecto de los planos coordenados xy, xz, yz, denotados por M_{xy}, M_{xz}, M_{yz}, son, respectivamente

$$M_{xy} = \iiint\limits_{\Omega} z\rho(x, y, z)dx\,dy\,dz,$$

$$M_{xz} = \iiint\limits_{\Omega} y\rho(x, y, z)dx\,dy\,dz$$

$$M_{yz} = \iiint\limits_{\Omega} x\rho(x, y, z)dx\,dy\,dz$$

y el centro de masa se localiza en el punto $(\bar{x}, \bar{y}, \bar{z})$ donde

$$\bar{x} = \frac{M_{yz}}{M} = \frac{\iiint\limits_{\Omega} x\rho(x, y, z)\,dx\,dy\,dz}{\iiint\limits_{\Omega} \rho(x, y, z)\,dx\,dy\,dz}$$

$$\bar{y} = \frac{M_{xz}}{M} = \frac{\iiint\limits_{\Omega} y\rho(x, y, z)dx\,dy\,dz}{\iiint\limits_{\Omega} \rho(x, y, z)dx\,dy\,dz}$$

$$\bar{z} = \frac{M_{xy}}{M} = \frac{\iiint\limits_{\Omega} z\rho(x, y, z)dx\,dy\,dz}{\iiint\limits_{\Omega} \rho(x, y, z)dx\,dy\,dz}$$

Si el cuerpo es homogéneo, lo cual significa que tiene densidad constante, digamos $\rho(x, y) = \rho_0$, entonces su masa total es

$$\begin{aligned} M &= \iiint\limits_{\Omega} \rho(x, y, z)\,dx\,dy\,dz = \rho_0 \iiint\limits_{\Omega} dx\,dy\,dz \\ &= (\text{ densidad })(\text{ volumen del cuerpo }) \end{aligned}$$

y las coordenadas de su centro de masa se escribirían como

$$\bar{x} = \frac{\iiint\limits_{\Omega} x\,dx\,dy\,dz}{\iiint\limits_{\Omega} dx\,dy\,dz}, \quad \bar{y} = \frac{\iiint\limits_{\Omega} y\,dx\,dy\,dz}{\iiint\limits_{\Omega} dx\,dy\,dz}, \quad \bar{z} = \frac{\iiint\limits_{\Omega} z\,dx\,dy\,dz}{\iiint\limits_{\Omega} dx\,dy\,dz}$$

Los momentos de inercia del cuerpo Ω respecto de los ejes coordenados x, y, z, denotados por I_X, I_Y e I_Z respectivamente, son

$$I_X = \iiint_{\Omega} (y^2 + z^2)\rho(x, y, z)\, dx\, dy\, dz$$

$$I_y = \iiint_{\Omega} (x^2 + z^2)\rho(x, y, z)\, dx\, dy\, dz$$

$$I_Z = \iiint_{\Omega} (x^2 + y^2)\rho(x, y, z)\, dx\, dy\, dz$$

y el momento de inercia respecto del origen (llamado "momento polar de inercia") es

$$I_0 = \iiint_{\Omega} (x^2 + y^2 + z^2)\rho(x, y, z)\, dx\, dy\, dz$$

Ejemplo 8. Hallemos el centro de masa del cuerpo homogéneo Ω limitado por el paraboloide $z = x^2 + y^2$ y el plano $z = 1$. Por consideraciones de simetría de Ω, podemos concluir desde el principio que el centro de masa procurado se encuentra sobre el eje z, es decir, que $\bar{x} = \bar{y} = 0$. Hallemos entonces $\bar{z}$. La proyección de Ω sobre el plano xy es $x^2 + y^2 \leq 1$, que se describe fácilmente en coordenadas cilíndricas. En estas coordenadas el paraboloide $z = x^2 + y^2$ se ve como $z = r^2$. La región Ω' que corresponde a Ω es entonces

$$\Omega' = \{(r, \theta, z) | 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi, r^2 \leq z \leq 1\}$$

Así, el volumen V de Ω es

$$V = \iiint_{\Omega'} r\, dr\, d\theta\, dz = \int_0^{2\pi} \left(\int_0^1 \left[\int_{r^2}^1 r\, dz \right] dr \right) d\theta$$

$$= 2\pi \int_0^1 (1 - r^2) r\, dr = \frac{\pi}{2}$$

La integral triple del numerador de $\bar{z}$ se calcula también en coordenadas cilíndricas:

$$\iiint_{\Omega} z\, dx\, dy\, dz = \int_0^{2\pi} \left(\int_0^1 \left[\int_{r^2}^1 zr\, dz \right] dr \right) d\theta$$

$$= \frac{1}{2} \int_0^{2\pi} \left(\int_0^1 (1 - r^4) r\, dr \right) d\theta = \pi \int_0^1 (r - r^5)\, dr = \frac{\pi}{3}$$

Entonces la coordenada $\bar{z}$ del centro de masa de Ω es

$$\bar{z} = \frac{\pi/3}{\pi/2} = \frac{2}{3}$$

Es decir, el centro de masa del cuerpo se encuentra en $(0, 0, 2/3)$. ■

Ejemplo 9. Se quiere calcular el momento de inercia de un cubo homogéneo de densidad ρ_0, de lado a, respecto de una de sus aristas. Sea entonces Ω el cubo

$$\Omega = \{(x, y, z)|0 \le x \le a, 0 \le y \le a, 0 \le z \le a\}$$

y calculemos su momento de inercia respecto de alguno (cualquiera) de los ejes coordenados, digamos el eje x. Se tiene entonces

$$\begin{aligned} I_X &= \iiint_{\Omega} (y^2 + z^2)\rho(x, y, z)\, dx\, dy\, dz = \rho_0 \int_0^a \int_0^a \int_0^a (x^2 + y^2)\, dx\, dy\, dz \\ &= \rho_0 \left[\int_0^a x^2\, dx \int_0^a dy \int_0^a dz + \int_0^a dx \int_0^a y^2\, dy \int_0^a dz \right] \\ &= \rho_0 \left[(a^3/3)(a)(a) + (a)(a^3/3)(a) \right] = \frac{2}{3}\rho_0 a^5 \end{aligned}$$

■

6.8.3 Valor medio de una función

Sea Ω una región acotada en $\mathbb{R}^3$. Sea $f: \Omega \to \mathbb{R}$ una función continua definida en Ω. El *valor medio de f en Ω*, denotado por $\bar{f}_\Omega$, se define como

$$\bar{f}_\Omega = \frac{\displaystyle\iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz}{\displaystyle\iiint_{\Omega} dx\, dy\, dz}$$

Es posible establecer un resultado análogo al teorema 6.5.1 de la sección 5 (subsección 6.5.4): el valor $\bar{f}_\Omega$ es tomado por la función f en un punto (a, b, c) de Ω (es decir, se da $(a, b, c) \in \Omega$ de modo que $\bar{f}_\Omega = f(a, b, c)$).

Ejemplo 10. Consideremos un cuerpo en $\mathbb{R}^3$ que ocupa la región Ω limitada por la superficie $(x^2 + y^2 + z^2)^2 = x^2 + y^2$. Sea $\rho: \Omega \to \mathbb{R}$, $\rho(x, y, z) = (x^2 + y^2 + z^2)^{3/2}$ la función que da la densidad $\rho(x, y, z)$ del cuerpo en el punto $(x, y, z) \in \Omega$. La densidad media del cuerpo en Ω será justamente el valor medio de la función ρ en Ω. Calculémosla. Pasando a las coordenadas esféricas, la ecuación de la superficie se transforma en $r = \operatorname{sen} \phi$, y la de la función densidad en $\rho(r \operatorname{sen} \phi \cos \theta, r \operatorname{sen} \phi \operatorname{sen} \theta, r \cos \theta) = r^3$. El denominador de $\bar{\rho}_\Omega$ es el volumen de Ω. Este es

$$\begin{aligned} \iiint_{\Omega} dx\, dy\, dz &= \int_0^{2\pi} \int_0^{\pi} \int_0^{\operatorname{sen} \phi} r^2 \operatorname{sen} \phi\, dr\, d\phi\, d\theta \\ &= \frac{1}{3} 2\pi \int_0^{\pi} \operatorname{sen}^4 \phi\, d\phi = \frac{1}{3} 2\pi \left(\frac{3\pi}{8} \right) = \frac{1}{4}\pi^2 \end{aligned}$$

Calculemos ahora el numerador de $\bar{\rho}_{\Omega}$. Se tiene

$$\iiint_{\Omega} \rho(x, y, z)\, dx\, dy\, dz = \int_0^{2\pi} \int_0^{\pi} \int_0^{\operatorname{sen}\phi} r^3 r^2 \operatorname{sen}\phi\, dr\, d\phi\, d\theta$$

$$= \frac{1}{6} 2\pi \int_0^{\pi} \operatorname{sen}^7 \phi\, d\phi = \frac{1}{6} 2\pi \left(\frac{32}{35}\right) = \frac{32\pi}{105}$$

Entonces la densidad media del cuerpo en Ω es

$$\bar{\rho}_{\Omega} = \frac{\displaystyle\iiint_{\Omega} \rho(x, y, z)\, dx\, dy\, dz}{\displaystyle\iiint_{\Omega} dx\, dy\, dz} = \frac{\frac{32\pi}{105}}{\frac{1}{4}\pi^2} = \frac{128}{105\pi}$$

■

Ejercicios (Capítulo 6, Sección 8)

En los ejercicios 1—10, calcule el volumen del cuerpo limitado por las superficies dadas.

1. $x^2 + y^2 = 1$, $x^2 + z^2 = 1$.

2. $x^2 + y^2 = 1$, $z = 4 - x^2 - y^2$, $z = 0$.

3. $z = x^2 + y^2$, $z = 2x^2 + 2y^2$, $z = 2$.

4. $z^2 = x^2 + y^2$, $z = 6 - x^2 - y^2$, $z \geq 0$.

5. $x^2 + y^2 + z^2 = 2z$, $x^2 + y^2 \leq z^2$.

6. $(x^2 + y^2 + z^2)^3 = 3xyz$

7. $x^2 + y^2 + z^2 = 1$, $x^2 + y^2 + z^2 = 4$, $x^2 + y^2 = z^2$, $z \geq 0$.

8. $\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$

9. $\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}$

10. $\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)^2 = x$

11. Considere la esfera $x^2 + y^2 + z^2 = a^2$. Tomando como vértice el punto $(0, 0, b)$, en donde $b > a > 0$, se traza un cono que "envuelve" tangencialmente a la esfera. Calcular el volumen del cuerpo comprendido entre el cono y la esfera, así como su área. (Nota: este problema puede resolverse de manera "elemental"; con las herramientas del cálculo en una variable. Ver también el ejercicio 31 de la sección 5 de este capítulo, y el ejercicio 45 de la sección 10 del capítulo 2).

12. Halle la masa total del cubo $\Omega = [-1, 1] \times [-1, 1] \times [-1, 1]$, sabiendo que su densidad en el punto $P = (x, y, z)$ es $\rho(x, y, z) = x + y + z$.

13. Halle la masa total de la esfera $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 1\}$, sabiendo que su densidad en el punto P es directamente proporcional a la distancia de P al origen, siendo igual a la unidad en los puntos de la frontera de Ω.

En los ejercicios 14—21, determine el centro de masa de los cuerpos homogéneos limitados por las superficies indicadas.

14. $x + y + z = 1, x = 0, y = 0, z = 0.$

15. $x^2 + y^2 + z^2 = 1, z \geq 0.$

16. $x^2 + y^2 + z^2 = 1, y \geq 0, z \geq 0.$

17. $x^2 + y^2 + z^2 = 1, x \geq 0, y \geq 0, z \geq 0.$

18. $z = y^2, z = 1, x = -1, x = 1.$

19. $2z = x^2 + y^2, x^2 + y^2 + z^2 = 3, z \geq 0.$

20. $z^2 = x^2 + y^2, x^2 + y^2 + z^2 = 1, z \geq 0.$

21. $z = (x^2 + y^2 + z^2)^2.$

22. Halle el momento de inercia de un paralelepípedo homogéneo de densidad ρ_0, de lados a, b, c, cuya masa total es M, con respecto a cada uno de sus lados.

23. Halle el momento de inercia de una esfera homogénea de densidad ρ_0, de radio R, y de masa total M, con respecto de uno de sus diámetros y con respecto de una tangente a ella.

24. Halle el momento de inercia de un elipsoide homogéneo de densidad ρ_0, de semiejes a, b, c, y masa total M, con respecto a cada uno de sus ejes.

En los ejercicios 25—29, determine el valor medio de la función $f(x, y, z)$ dada en la región indicada.

25. $f(x, y, z) = x + y + z$, en $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 1\}$.

26. $f(x, y, z) = xyz$, en $\Omega = \{(x, y, z)|x^2 + y^2 + z^2 \leq 1\}$.

27. $f(x, y, z) = x$, en $\Omega = [0, 1] \times [0, 1] \times [0, 1]$.

28. $f(x, y, z) = x^2 + y^2 + z^2$, en $\Omega = \{(x, y, z)|x^2 + y^2 \leq 1, -1 \leq z \leq 1\}$.

29. $f(x, y, z) = x^2 + y^2 + z$, en $\Omega = \{(x, y, z)|1 \leq x^2 + y^2 + z^2 \leq 4\}$.

30. Determine la densidad media del cuerpo del ejercicio 12.

31. Halle la densidad media del cuerpo del ejercicio 13.

32. Calcule el valor promedio de la suma de tres números, si la suma de sus cuadrados es siempre no mayor a la unidad. (Ver ejercicio 25).

33. Calcule el valor promedio del producto de tres números, si la suma de sus cuadrados es siempre no mayor a la unidad. (Ver ejercicio 26).

6.9 Integrales N-múltiples

En esta última sección del capítulo presentamos las generalizaciones naturales de los conceptos estudiados en las secciones anteriores sobre integrales de funciones de dos y tres variables, a integrales de funciones de n variables. Esta es, por decirlo así, la sección que resume las ideas importantes del capítulo, expuestas ahora, en general, para funciones con n variables.

Como se ha hecho anteriormente para $n = 2$ y $n = 3$, el primer paso que daremos será definir integral n-múltiple para funciones escalonadas de n variables $f: Q \subset \mathbb{R}^n \to \mathbb{R}$ definidas en productos cartesianos de n intervalos de $\mathbb{R}$, $Q = I_1 \times I_2 \times \ldots \times I_n$.

Sean $I_1 = [a_1, b_1]$, $I_2 = [a_2, b_2], \ldots, I_n = [a_n, b_n]$ n intervalos de $\mathbb{R}$. Llamaremos al producto cartesiano

$$Q = I_1 \times I_2 \times \ldots \times I_n = \{(x_1, x_2, \ldots, x_n) | x_i \in I_i,\ i = 1, 2, \ldots, n\}$$

rectángulo en $\mathbb{R}^n$. Una partición P de Q es un producto cartesiano del tipo $P = P_1 \times P_2 \times \ldots \times P_n$ donde P_i es una partición del intervalo I_i, $i = 1, 2, \ldots, n$. Digamos que P_i divide al intervalo I_i en m_i subintervalos. Entonces P divide a Q en $m_1 m_2 \ldots m_n$ subrectángulos en $\mathbb{R}^n$ y escribiremos

$$Q_{i_1 i_2 \ldots i_n} = [x_{i_1 - 1}, x_{i_1}] \times [x_{i_2 - 1}, x_{i_2}] \times \ldots \times [x_{i_n - 1}, x_{i_n}]$$

para denotar al $i_1 i_2 \ldots i_n$-ésimo subrectángulo de la partición P de Q, $i_1 = 1, 2, \ldots, m_1$, $i_2 = 1, 2, \ldots, m_2, \ldots, i_n = 1, 2, \ldots, m_n$.

Se dice que la función $f: Q \subset \mathbb{R}^n \to \mathbb{R}$ definida en el rectángulo Q en $\mathbb{R}^n$ es una *función escalonada*, si se da una partición P de Q, digamos que en $m_1 m_2 \ldots m_n$ subrectángulos $Q_{i_1 i_2 \ldots i_n}$, tal que f restringida a cada uno de estos subrectángulos es constante. Digamos que $f|_{Q_{i_1 i_2 \ldots i_n}} = c_{i_1 i_2 \ldots i_n} \in \mathbb{R}$. Se define la *integral n-múltiple* de la función f de las variables $x_1, x_2, \ldots, x_n$, sobre el rectángulo Q, denotada por

$$\iint_Q \cdots \int f \quad \text{o} \quad \iint_Q \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$

como

$$\begin{aligned}&\iint_Q \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n \\ &= \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_n=1}^{m_n} c_{i_1 i_2 \ldots i_n} (x_{i_1} - x_{i_1 - 1})(x_{i_2} - x_{i_2 - 1}) \ldots (x_{i_n} - x_{i_n - 1})\end{aligned}$$

Al producto $(x_{i_1} - x_{i_1 - 1})(x_{i_2} - x_{i_2 - 1}) \cdots (x_{i_n} - x_{i_n - 1})$ lo llamaremos "volumen del $i_1 i_2 \ldots i_n$-ésimo subrectángulo $Q_{i_1 i_2 \ldots i_n} = [x_{i_1 - 1}, x_{i_1}] \times [x_{i_2 - 1}, x_{i_2}] \times \ldots \times [x_{i_n - 1}, x_{i_n}]$ de la partición P de Q". Esta terminología se justifica más adelante (ver ejemplo 1). Obsérvese entonces que podemos escribir la integral n-múltiple de f sobre Q como

$$\begin{aligned}&\iint_Q \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n \\ &= \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \ldots \sum_{i_n=1}^{m_n} (\text{valor de } f \text{ en } Q_{i_1 i_2 \ldots i_n}) (\text{ volumen de } Q_{i_1 i_2 \ldots i_n})\end{aligned}$$

Una vez establecido el concepto de integral n-múltiple para funciones escalonadas, éste se establece ahora para funciones acotadas (se dice que la función $f: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto Ω de $\mathbb{R}^n$ es acotada si ocurre un $M > 0$ de modo que $|f(x_1, x_2, \ldots, x_n)| < M$ para todo $(x_1, x_2, \ldots, x_n) \in \Omega$).

Definición. Sea $f: Q \subset \mathbb{R}^n \to \mathbb{R}$ una función acotada definida en el rectángulo Q de $\mathbb{R}^n$. Si ocurre un único número real I de modo que

$$\iint_Q \cdots \int \phi \leq I \leq \iint_Q \cdots \int \psi$$

para todas las funciones escalonadas ϕ, $\psi: Q \subset \mathbb{R}^n \to \mathbb{R}$, que $\phi(x_1, x_2, \ldots, x_n) \leq f(x_1, x_2, \ldots, x_n) \leq \psi(x_1, x_2, \ldots, x_n)$, $\forall (x_1, x_2, \ldots, x_n) \in Q$ se dice que la función f es *integrable* sobre Q, y al número I se le llama integral n-múltiple de f sobre Q y se escribe

$$I = \iint_Q \cdots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$$ ■

Por supuesto, las funciones continuas están dentro del tipo de funciones de las que se habla en la definición anterior.

Teorema 6.9.1 Sea $f: Q \subset \mathbb{R}^n \to \mathbb{R}$ una función continua definida en el rectángulo Q de $\mathbb{R}^n$. Entonces f es integrable en Q. ■

Podemos ser más precisos en la conclusión del teorema anterior e indicar cómo se puede calcular la integral de f sobre $Q = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$. Se tiene de hecho que

$$\iint_Q \cdots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$$
$$= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\cdots \int_{a_n}^{b_n} f(x_1, x_2, \ldots, x_n) dx_n \ldots \right) dx_2 \right) dx_1$$

En la expresión anterior es posible hacer intercambios en el orden de integración, como se señaló en el caso de las integrales dobles y triples. Incluso si la función f no tiene demasiadas discontinuidades, ésta sigue siendo integrable, como lo establece el siguiente teorema.

Teorema 6.9.2 Sea $f: Q \subset \mathbb{R}^n \to \mathbb{R}$ una función definida y acotada en el rectángulo Q de $\mathbb{R}^n$. Si el conjunto de discontinuidades de f en Q tiene medida cero, entonces f es integrable en Q. ■

Un conjunto $C \subset \mathbb{R}^n$ tiene medida cero si se dan rectángulos Q_i de $\mathbb{R}^n$, $i = 1, 2, \ldots, n$ cuya unión contiene a C y cuya suma de volúmenes es menor que cualquier número $\epsilon > 0$ previamente dado. Un ejemplo importante de conjunto de medida cero en $\mathbb{R}^3$ es el dado por la gráfica de una función continua $\phi: K \subset \mathbb{R}^{n-1} \to \mathbb{R}$ definida en un conjunto compacto (cerrado y acotado) de $\mathbb{R}^{n-1}$. Este ejemplo es el análogo del teorema 6.2.3.

Para extender el concepto de integral n-múltiple de funciones continuas de n variables a regiones más generales en $\mathbb{R}^n$, que denotamos por Ω, procedemos de una manera análoga a la expuesta en los

casos de las integrales dobles y triples. Una región típica Ω sobre la que efectuaremos una integral n-múltiple, es la que forman los puntos $(x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$, que

$$a \leq x_1 \leq b \tag{1}$$
$$f_1^1(x_1) \leq x_2 \leq f_2^1(x_1) \tag{2}$$
$$f_1^2(x_1, x_2) \leq x_3 \leq f_2^2(x_1, x_2) \tag{3}$$
$$\vdots$$
$$f_1^{n-1}(x_1, x_2, \ldots, x_{n-1}) \leq x_n \leq f_2^{n-1}(x_1, x_2, \ldots, x_{n-1}) \tag{n}$$

donde f^i es una función real continua de las variables $x_1, x_2, \ldots, x_i$ definida en $(x_1, x_2, \ldots, x_i)$ de modo que x_j satisface la desigualdad $1 \leq j \leq i$. Es claro que este tipo de regiones en $\mathbb{R}^n$ son acotadas. Es decir, se puede conseguir un rectángulo Q en $\mathbb{R}^n$ que contenga a la región. Para introducir el concepto de integral n-múltiple de una función continua $f(x_1, x_2, \ldots, x_n)$ sobre la región Ω, tomamos el rectángulo Q en $\mathbb{R}^n$ que contiene a Ω y definimos en Q la función $\tilde{f}(x_1, x_2, \ldots, x_n)$ como

$$\tilde{f}(x_1, x_2, \ldots, x_n) = \begin{cases} f(x_1, x_2, \ldots, x_n) & \text{si } (x_1, x_2, \ldots, x_n) \in \Omega \\ 0 & \text{si } (x_1, x_2, \ldots, x_n) \in Q \backslash \Omega \end{cases}$$

Se puede demostrar que, siendo la función f continua en Ω, la función $\tilde{f}(x_1, x_2, \ldots, x_n)$ es integrable sobre Q (sus discontinuidades forman un conjunto de medida cero, en el que están incluídos los puntos de las gráficas de las funciones $x_n = f_1^{n-1}(x_1, x_2, \ldots, x_{n-1})$ y $x_n = f_2^{n-1}(x_1, x_2, \ldots, x_{n-1})$). Así, la integral de f sobre Ω se define como la integral de $\tilde{f}$ sobre Q. Es decir

$$\iint_{\Omega} \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n = \iint_{Q} \cdots \int \tilde{f}(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$

En tal caso, se puede demostrar que la integral n-múltiple de la función f se calcula como

$$\iint_{\Omega} \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$
$$= \int_a^b \int_{f_1^1(x_1)}^{f_2^1(x_1)} \int_{f_1^2(x_1,x_2)}^{f_2^2(x_1,x_2)} \cdots \int_{f_1^{n-1}(x_1,x_2,\ldots,x_{n-1})}^{f_2^{n-1}(x_1,x_2,\ldots,x_{n-1})} f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$

donde la primera integral (la más interna) se calcula respecto de la variable x_n (con límites de integración $f_1^{n-1}(x_1, x_2, \ldots, x_{n-1})$ y $f_2^{n-1}(x_1, x_2, \ldots, x_{n-1})$) y la última se calcula respecto de la variable x_1 (con límites de integración a y b).

Al calcular la integral n-múltiple de la función $f(x_1, x_2, \ldots, x_n)$ sobre la región Ω de $\mathbb{R}^n$, diremos que se está calculando un "hipervolumen" (en $\mathbb{R}^{n+1}$) debajo de la gráfica de la función $x_{n+1} = f(x_1, x_2, \ldots, x_n)$ —llamada "hipersuperficie"— sobre Ω. En particular, la integral de la función $f(x_1, x_2, \ldots, x_n) = 1$ sobre Ω es el "volumen" de la región Ω. Es decir

$$\iint_{\Omega} \cdots \int dx_1 dx_2 \ldots dx_n = \text{volumen de } \Omega$$

Ejemplo 1. Consideremos el rectángulo $Q = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$. Según lo antes expuesto, el volumen de este rectángulo es

$$\iint_Q \cdots \int dx_1 dx_2 \ldots dx_n = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\ldots \int_{a_n}^{b_n} dx_n \ldots \right) dx_2 \right) dx_1$$
$$= (b_1 - a_1)(b_2 - a_2) \ldots (b_n - a_n)$$

■

Ejemplo 2. (Volumen de un $(N+1)$-edro en $\mathbb{R}^n$). Se quiere calcular el volumen de la región Ω en $\mathbb{R}^n$ limitada por

$$x_1 = 0, x_2 = 0, \ldots, x_n = 0, \quad \frac{x_1}{a_1} + \frac{x_2}{a_2} + \cdots + \frac{x_n}{a_n} = 1$$

donde $a_1, a_2, \ldots, a_n$ son números reales positivos dados. Este ejemplo generaliza al presentado en el ejemplo 1 de la sección 5, en el que se calculó el volumen del tetraedro en $\mathbb{R}^3$ limitado por los planos coordenados y por el plano $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ (y que resultó ser $\frac{1}{6}abc$). Nuestro problema se puede ver como el cálculo del volumen de la región Ω limitada por

$$0 \le x_1 \le a_1$$
$$0 \le x_2 \le a_2\left(1 - \frac{x_1}{a_1}\right)$$
$$0 \le x_3 \le a_3\left(1 - \frac{x_1}{a_1} - \frac{x_2}{a_2}\right)$$
$$\vdots$$
$$0 \le x_n \le a_n\left(1 - \frac{x_1}{a_1} - \frac{x_2}{a_2} - \cdots - \frac{x_{n-1}}{a_{n-1}}\right)$$

de modo que el volumen buscado es

$$\iint_\Omega \cdots \int dx_1 dx_2 \ldots dx_n$$
$$= \int_0^{a_1} \int_0^{a_2(1-x_1/a_1)} \cdots \int_0^{a_n(1-x_1/a_1-\cdots-x_{n-1}/a_{n-1})} dx_n \ldots dx_2 dx_1$$
$$\int_0^{a_1} \cdots \int_0^{a_{n-1}(1-x_1/a_1-\cdots-x_{n-2}/a_{n-2})} a_n\left(1 - \frac{x_1}{a_1} - \frac{x_2}{a_2} - \cdots - \frac{x_{n-1}}{a_{n-1}}\right) dx_{n-1} \ldots dx_1$$
$$= \frac{1}{2}a_n a_{n-1} \int_0^{a_1} \cdots \int_0^{a_{n-2}(1-x_1/a_1-\cdots-x_{n-3}/a_{n-3})} \left(1 - \frac{x_1}{a_1} - \cdots - \frac{x_{n-2}}{a_{n-2}}\right)^2 dx_{n-2} \ldots dx_1$$
$$= \frac{1}{6}a_n a_{n-1} a_{n-2} \int_0^{a_1} \cdots \int_0^{a_{n-3}(1-x_1/a_1-\cdots-x_{n-4}/a_{n-4})} \left(1 - \frac{x_1}{a_1} - \cdots - \frac{x_{n-3}}{a_{n-3}}\right)^3 dx_{n-3} \ldots dx_1$$

Continuando con estos cálculos se obtiene finalmente el valor

$$\frac{a_1 a_2 \ldots a_n}{n!}$$

Esta integral se puede calcular con un cambio de variables, tema que estudiaremos a continuación. En el ejemplo siguiente la recalcularemos con esta nueva técnica. ■

Supongamos que en la integral

$$\iint_{\Omega} \cdots \int f(x_1, x_2, \cdots, x_n)\, dx_1 dx_2 \ldots dx_n$$

(donde f es una función continua en Ω) se quieren cambiar las n variables $x_1, x_2, \ldots, x_n$ por otras n variables $u_1, u_2, \ldots, u_n$, según las fórmulas (de cambio de variables)

$$\begin{aligned} x_1 &= x_1(u_1, u_2, \ldots, u_n) \\ x_2 &= x_2(u_1, u_2, \ldots, u_n) \\ &\vdots \\ x_n &= x_n(u_1, u_2, \ldots, u_n) \end{aligned}$$

o bien, según la función $F: \Omega' \to \mathbb{R}^n$ (llamada función de transformación de coordenadas)

$$\begin{aligned} F(u_1, u_2, \ldots, u_n) &= (x_1, x_2, \ldots, x_n) \\ &= \big(x_1(u_1, u_2, \ldots, u_n), x_2(u_1, u_2, \ldots, u_n), \ldots, x_n(u_1, u_2, \ldots, u_n)\big) \end{aligned}$$

que manda los puntos $(u_1, u_2, \ldots, u_n)$ de una región Ω' de $\mathbb{R}^n$ en los puntos $(x_1, x_2, \ldots, x_n)$ de la región Ω de $\mathbb{R}^n$. Si esta función es de clase $\mathscr{C}^1$, inyectiva, y si su determinante jacobiano $\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2 \ldots u_n)}$ es distinto de cero en Ω', (o bien, es distinto de cero excepto en un subconjunto de medida cero de Ω') se tiene la fórmula del cambio de variables en integrales n-múltiples

$$\begin{aligned} &\iint_{\Omega} \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n \\ &= \iint_{\Omega'} \cdots \int f(F(u_1, u_2, \ldots, u_n)) \left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} \right| du_1 du_2 \ldots du_n \end{aligned}$$

En este caso el jacobiano $\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)}$ es el determinante de la matriz $n \times n$ cuyo elemento de la i-ésima línea y j-ésima columna es $\frac{\partial x_i}{\partial u_j}$, $i, j = 1, 2, \ldots, n$. Como se sabe del capítulo 3 (ver ejercicio 18 de la sección 6 del capítulo 3), se tiene

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} \frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} = 1$$

(algunas veces es más sencillo calcular el segundo jacobiano).

En particular, al calcular el volumen de una región Ω de $\mathbb{R}^n$ cambiando las variables x_i por las variables u_i según se ha discutido previamente, se tiene

$$\begin{aligned}\text{volumen de } \Omega &= \iint_{\Omega}\cdots\int dx_1 dx_2 \dots dx_n \\ &= \iint_{\Omega'}\cdots\int \left|\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(u_1, u_2, \dots, u_n)}\right| du_1 du_2 \dots du_n\end{aligned}$$

Ejemplo 3. Regresemos a recalcular la integral del ejemplo 2

$$\int_0^{a_1}\int_0^{a_2(1-x_1/a_1)}\cdots\int_0^{a_n(1-x_1/a_1-\cdots-x_{n-1}/a_{n-1})} dx_1 dx_2 \dots dx_n$$

Introduzcamos las nuevas variables $u_1, u_2, \dots, u_n$ según las fórmulas

$$u_1 = \frac{x_1}{a_1}, \quad u_2 = \frac{\frac{x_2}{a_2}}{1-\frac{x_1}{a_1}}, \quad u_3 = \frac{\frac{x_3}{a_3}}{1-\frac{x_1}{a_1}-\frac{x_2}{a_2}}, \dots, u_n = \frac{\frac{x_n}{a_n}}{1-\frac{x_1}{a_1}-\cdots-\frac{x_{n-1}}{a_{n-1}}},$$

Nótese entonces que la región Ω corresponde a la región Ω' en la que

$$0 \leq u_i \leq 1, \quad i = 1, 2, \dots, n$$

Es decir, Ω' es el rectángulo en $\mathbb{R}^n$, $\Omega' = [0, 1] \times [0, 1] \times \dots \times [0, 1]$ (denotado como $[0, 1]^n$). Calculemos el jacobiano de la función de transformación de coordenadas $F(u_1, u_2, \dots, u_n) = (x_1, x_2, \dots, x_n)$. En este caso es fácil hacer explícitas las funciones coordenadas de F, despejando las variables x_i en términos de las variables u_i de las expresiones anteriormente establecidas. Se tiene de hecho que

$$\begin{aligned}x_1 = a_1 u_1, \quad x_2 = a_2 u_2(1-u_1), \quad x_3 = a_3 u_3(1-u_1)(1-u_2), \dots, \\ x_n = a_n u_n(1-u_1)(1-u_2)\cdots(1-u_{n-1})\end{aligned}$$

Nótese entonces que el jacobiano $\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(u_1, u_2, \dots, u_n)}$ es el determinante de una matriz triangular inferior, pues $\frac{\partial x_i}{\partial u_j} = 0$ para $i < j$. Por tanto, sólo basta calcular los elementos de la diagonal de esta matriz, pues el producto de éstos es el determinante procurado. Estos elementos son $\frac{\partial x_i}{\partial u_i}$, $i = 1, 2, \dots, n$. Es decir que

$$\begin{aligned}\frac{\partial(x_1, x_2, \dots, x_n)}{\partial(u_1, u_2, \dots, u_n)} &= \left(\frac{\partial x_1}{\partial u_1}\right)\left(\frac{\partial x_2}{\partial u_2}\right)\cdots\left(\frac{\partial x_n}{\partial u_n}\right) \\ &= a_1 a_2 \dots a_n (1-u_1)^{n-1}(1-u_2)^{n-2}\dots(1-u_{n-1})\end{aligned}$$

Entonces el volumen procurado es

$$\begin{aligned}
&\iint_{[0,1]^n}\cdots\int a_1a_2\ldots a_n(1-u_1)^{n-1}(1-u_2)^{n-2}\ldots(1-u_{n-1})\,du_1du_2\ldots du_n\\
&= a_1a_2\ldots a_n\int_0^1(1-u_1)^{n-1}\,du_1\int_0^1(1-u_2)^{n-2}\,du_2\ldots\int_0^1(1-u_{n-1})\,du_{n-1}\int_0^1 du_n\\
&= a_1a_2\ldots a_n\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\cdots\left(\frac{1}{2}\right)(1)=\frac{a_1a_2\ldots a_n}{n!}
\end{aligned}$$

Nótese que para $n=2$, este resultado da el área de un triángulo de base a_1 y altura a_2, para $n=3$, da el volumen de un tetraedro, como se calculó en el ejemplo 1 de la sección 5. ■

Ejemplo 4. (Volumen de un paralelepípedo en $\mathbb{R}^n$). Para generalizar el ejercicio 2 de la sección 8, en el que se calculó el volumen de un paralelepípedo en $\mathbb{R}^3$ limitado por 3 parejas de planos paralelos, consideremos ahora los $2n$ planos en $\mathbb{R}^n$

$$a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_{11},\quad a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_{12}\qquad(\Pi_1)$$
$$a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_{21},\quad a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_{22}\qquad(\Pi_2)$$
$$\ldots$$
$$a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_{n1},\quad a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_{n2}\qquad(\Pi_n)$$

donde $b_{i1}<b_{i2}$, $i=1,2,\ldots,n$, y

$$\Delta=\det\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \vdots & \vdots & \vdots & \\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{bmatrix}\neq 0$$

Se quiere calcular el volumen del cuerpo Ω en $\mathbb{R}^n$ limitado por estos planos. Es decir, se quiere calcular la integral

$$\iint_{\Omega}\cdots\int dx_1dx_2\ldots dx_n$$

Introduzcamos las variables $u_1, u_2, \ldots, u_n$ definidas por

$$\begin{aligned}
u_1&=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n\\
u_2&=a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n\\
&\ldots\\
u_n&=a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n
\end{aligned}$$

La región Ω' que corresponde a Ω es el rectángulo en $\mathbb{R}^n$

$$\Omega'=[b_{11},b_{12}]\times[b_{21},b_{22}]\times\ldots\times[b_{n1},b_{n2}]$$

Además, es claro que

$$\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} = \Delta$$

de modo que

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} = \frac{1}{\Delta}$$

y entonces el volumen procurado es

$$\iint_{\Omega} \cdots \int dx_1 dx_2 \ldots dx_n = \iint_{\Omega'} \cdots \int \frac{1}{\Delta} du_1 du_2 \ldots du_n$$
$$= \frac{1}{\Delta}(b_{12} - b_{11})(b_{22} - b_{21}) \cdots (b_{n2} - b_{n1})$$

(aplicando de modo directo el resultado del ejemplo 1). ■

Un cambio de variables importante, que generaliza a las coordenadas esféricas estudiadas con las integrales triples, es el siguiente: introducimos las nuevas n variables r, ϕ_1, $\phi_2, \ldots, \phi_{n-1}$ según las fórmulas

$$\begin{aligned}
x_1 &= r \operatorname{sen} \phi_1 \operatorname{sen} \phi_2 \cdots \operatorname{sen} \phi_{n-3} \operatorname{sen} \phi_{n-2} \cos \phi_{n-1} \\
x_2 &= r \operatorname{sen} \phi_1 \operatorname{sen} \phi_2 \cdots \operatorname{sen} \phi_{n-3} \operatorname{sen} \phi_{n-2} \operatorname{sen} \phi_{n-1} \\
x_3 &= r \operatorname{sen} \phi_1 \operatorname{sen} \phi_2 \cdots \operatorname{sen} \phi_{n-3} \cos \phi_{n-2} \\
x_4 &= r \operatorname{sen} \phi_1 \operatorname{sen} \phi_2 \cdots \cos \phi_{n-3} \\
&\vdots \\
x_{n-1} &= r \operatorname{sen} \phi_1 \cos \phi_2 \\
x_n &= r \cos \phi_1
\end{aligned}$$

donde $r \geq 0$, $0 \leq \phi_i \leq \pi$, $i = 1, 2, \ldots, n-2$, $0 \leq \phi_{n-1} \leq 2\pi$. Se puede demostrar que el valor absoluto del jacobiano de la transformación es

$$\left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(r, \phi_1, \ldots, \phi_{n-1})} \right| = r^{n-1} \operatorname{sen}^{n-2} \phi_1 \operatorname{sen}^{n-3} \phi_2 \cdots \operatorname{sen} \phi_{n-2}$$

Entonces, al calcular el volumen de un cuerpo Ω en $\mathbb{R}^n$ con estas nuevas variables, debemos calcular

$$\iint_{\Omega'} \cdots \int r^{n-1} \operatorname{sen}^{n-2} \phi_1 \operatorname{sen}^{n-3} \phi_2 \cdots \operatorname{sen} \phi_{n-2} \, dr d\phi_1 d\phi_2 \ldots d\phi_{n-1}$$

donde Ω' es la región que corresponde a Ω en estas nuevas coordenadas. Terminamos esta sección con dos ejemplos interesantes en los que se usa el cambio anterior de variables.

Ejemplo 5. (Volumen de una esfera en $\mathbb{R}^n$). Se trata de calcular el volumen contenido por una esfera en $\mathbb{R}^n$ de radio R

$$x_1^2 + x_2^2 + \cdots + x_n^2 = R^2 \tag{1}$$

En principio observamos que el problema se reduce fácilmente al cálculo del volumen de una esfera de radio 1 (llamada esfera unitaria). En efecto, si introducimos las variables $y_1, y_2, \ldots, y_n$ de modo que $x_i = Ry_i$, $i = 1, 2, \ldots, n$, la ecuación (1) se transforma en la ecuación de la esfera unitaria

$$y_1^2 + y_2^2 + \cdots + y_n^2 = 1$$

El jacobiano de la transformación $F(y_1, y_2, \ldots, y_n) = (x_1, x_2, \ldots, x_n)$ es el determinante de la matriz diagonal de orden n $\text{diag}\left(\frac{\partial x_i}{\partial y_i}\right) = \text{diag}(R, R, \ldots, R) = R^n$, de modo que

$$\text{Volumen contenido por la esfera de radio } R = \iint\limits_{x_1^2 \cdots + x_n^2 \leq R^2} \cdots \int dx_1 dx_2 \ldots dx_n$$

$$= \iint\limits_{y_1^2 + \cdots + y_n^2 \leq 1} \cdots \int R^n \, dy_1 dy_2 \ldots dy_n = R^n \iint\limits_{y_1^2 + \cdots + y_n^2 \leq 1} \cdots \int dy_1 dy_2 \ldots dy_n$$

$$= R^n (\text{ Volumen de la esfera unitaria })$$

Denotaremos por S_n a la esfera unitaria en $\mathbb{R}^n$, y por $V(S_n)$ al volumen que ésta contiene. Entonces

$$V(S_n) = \iint\limits_{x_1^2 + \cdots + x_n^2 \leq 1} \cdots \int dx_1 dx_2 \ldots dx_n$$

Para calcular esta integral n-múltiple introducimos las variables $r, \phi_1, \phi_2, \ldots, \phi_{n-1}$ ya presentadas. Nótese que en estas nuevas variables la ecuación de la esfera S_n se transforma en $r = 1$. Así, el problema se reduce a calcular la integral

$$\iint\limits_{\Omega'} \cdots \int r^{n-1} \operatorname{sen}^{n-2} \phi_1 \operatorname{sen}^{n-3} \phi_2 \cdots \operatorname{sen} \phi_{n-2} \, dr d\phi_1 d\phi_2 \ldots d\phi_{n-1}$$

donde la nueva región Ω' es

$$\Omega' = \{(r, \phi_1, \phi_2, \ldots, \phi_{n-1}) | 0 \leq r \leq 1, \ 0 \leq \phi_{n-1} \leq 2\pi, \ 0 \leq \phi_i \leq \pi, \ i = 1, \ldots, n-2\}$$

Entonces $V(S_n)$ es igual a

$$\int_0^\pi \operatorname{sen}^{n-2} \phi_1 \, d\phi_1 \int_0^\pi \operatorname{sen}^{n-3} \phi_2 \, d\phi_2 \ldots \int_0^\pi \operatorname{sen} \phi_{n-2} \, d\phi_{n-2} \int_0^{2\pi} d\phi_{n-1} \int_0^1 r^{n-1} \, dr$$

y obtenemos así la siguiente fórmula explícita para el volumen de la esfera unitaria en $\mathbb{R}^n$ $(n \geq 3)$

$$V(S_n) = \frac{1}{n}(2\pi) \int_0^\pi \operatorname{sen}^{n-2} \phi_1 \, d\phi_1 \int_0^\pi \operatorname{sen}^{n-3} \phi_2 \, d\phi_2 \ldots \int_0^\pi \operatorname{sen} \phi_{n-2} \, d\phi_{n-2}$$

o bien

$$\boxed{V(S_n) = \frac{1}{n}(2\pi) \prod_{k=1}^{n-2} \int_0^\pi \operatorname{sen}^k \phi \, d\phi}$$

Nótese que si calculáramos el volumen de la esfera S_{n-1}, obtendríamos una expresión completamente análoga a la anterior, a saber

$$V(S_{n-1}) = \frac{1}{n-1}(2\pi)\int_0^\pi \operatorname{sen}^{n-3}\phi_1\,d\phi_1 \int_0^\pi \operatorname{sen}^{n-4}\phi_2\,d\phi_2 \ldots \int_0^\pi \operatorname{sen}\phi_{n-3}\,d\phi_{n-3}$$

o bien, puesto que se trata de un producto de integrales simples en las que los nombres de las variables de integración son irrelevantes,

$$V(S_{n-1}) = \frac{1}{n-1}(2\pi)\int_0^\pi \operatorname{sen}^{n-3}\phi_2\,d\phi_2 \ldots \int_0^\pi \operatorname{sen}\phi_{n-2}\,d\phi_{n-2}$$

de donde

$$\int_0^\pi \operatorname{sen}^{n-3}\phi_2\,d\phi_2 \ldots \int_0^\pi \operatorname{sen}\phi_{n-2}\,d\phi_{n-2} = \frac{n-1}{2\pi}V(S_{n-1}) \qquad (*)$$

por lo que el volumen de la esfera S_n se puede escribir como

$$V(S_n) = \frac{1}{n}(2\pi)\int_0^\pi \operatorname{sen}^{n-2}\phi_1\,d\phi_1 \frac{n-1}{2\pi}V(S_{n-1})$$

y así se obtiene la siguiente fórmula recursiva (para $n \geq 2$) del volumen de la esfera S_n

$$\boxed{V(S_n) = \frac{n-1}{n}V(S_{n-1})\int_0^\pi \operatorname{sen}^{n-2}\phi\,d\phi}$$

en la que se da el volumen de la esfera unitaria en $\mathbb{R}^n$ en términos del volumen de la esfera unitaria en $\mathbb{R}^{n-1}$. Partiendo de que la esfera S_1 en $\mathbb{R}$, (cuya ecuación es $x^2 = 1$, la cual determina entonces el intervalo $[-1, 1]$) tiene por "volumen" (que en ese caso se llama "longitud") $V(S_1) = 2$, resulta que

$$V(S_2) = \frac{1}{2}V(S_1)\int_0^\pi d\phi = \pi$$

(y entonces el volumen —que en este caso se llama "área"— de la esfera en $\mathbb{R}^2$ de radio R es πR^2). Del mismo modo

$$V(S_3) = \frac{2}{3}V(S_2)\int_0^\pi \operatorname{sen}\phi\,d\phi = \frac{2}{3}(\pi)(2) = \frac{4}{3}\pi$$

(y entonces el volumen de la esfera en $\mathbb{R}^3$ de radio R es $\frac{4}{3}\pi R^3$). El volumen de la esfera unitaria en $\mathbb{R}^4$ es

$$V(S_4) = \frac{3}{4}V(S_3)\int_0^\pi \operatorname{sen}^2\phi\,d\phi = \frac{3}{4}\left(\frac{4}{3}\pi\right)\left(\frac{\pi}{2}\right) = \frac{\pi^2}{2}$$

y el de la esfera unitaria en $\mathbb{R}^5$ es

$$V(S_5) = \frac{4}{5}V(S_4)\int_0^\pi \operatorname{sen}^3\phi\,d\phi = \frac{4}{5}\left(\frac{\pi^2}{2}\right)\left(\frac{4}{3}\right) = \frac{8\pi^2}{15}$$

etc. ■

Ejemplo 6. (Volumen del cono en $\mathbb{R}^n$). Consideremos la región Ω en $\mathbb{R}^3$ limitada por las superficies

$$x^2 + y^2 = \frac{R^2}{H^2}z^2, \quad z = H$$

donde R y H son números positivos dados. La primera de estas ecuaciones representa un cono (con dos hojas y vértice en el origen) y la segunda es un plano paralelo al plano xy. Así, el cuerpo limitado por dichas superficies es un cono circular recto de radio base R y altura H, cuyo volumen sabemos que es $\frac{1}{3}\pi R^2 H$. Este resultado se puede obtener haciendo la integral triple de la función 1 sobre Ω, la cual se puede calcular fácilmente en coordenadas esféricas $x = r \operatorname{sen} \phi \cos\theta$, $y = r \operatorname{sen}\phi \operatorname{sen}\theta$, $z = r\cos\phi$. En estas coordenadas la ecuación del cono toma la forma $\tan^2\phi = \frac{R^2}{H^2}$, o bien, $\phi = \arctan(R/H)$, y la ecuación del plano se ve como $r = H/\cos\phi$. Entonces el volumen buscado es

$$\begin{aligned}
&\int_0^{2\pi}\int_0^{\arctan(R/H)}\int_0^{H/\cos\phi} r^2 \operatorname{sen}\phi \, dr\, d\phi\, d\theta \\
&= \frac{2\pi}{3}H^3 \int_0^{\arctan(R/H)} \frac{\operatorname{sen}\phi}{\cos^3\phi}\, d\phi = \frac{\pi}{3}H^3\left(\sec^2\phi\right)_0^{\arctan(R/H)} \\
&= \frac{\pi}{3}H^3\left(\frac{R^2+H^2}{H^2} - 1\right) = \frac{1}{3}\pi R^2 H
\end{aligned}$$

En este ejercicio generalizamos estos cálculos al espacio $\mathbb{R}^n$. Consideramos el "cono circular recto" en $\mathbb{R}^n$

$$x_1^2 + x_2^2 + \cdots + x_{n-1}^2 = \frac{R^2}{H^2}x_n^2, \quad x_n = H$$

e introducimos las variables $r, \phi_1, \phi_2, \ldots, \phi_{n-1}$ según las ecuaciones presentadas antes del ejemplo 5. Obsérvese que

$$x_1^2 + x_2^2 + \cdots + x_{n-1}^2 = r^2 \operatorname{sen}^2\phi_1$$

de modo que la ecuación $x_1^2 + x_2^2 + \cdots + x_{n-1}^2 = \frac{R^2}{H^2}x_n^2$ se ve en estas nuevas coordenadas se ve como

$$r^2\operatorname{sen}^2\phi_1 = \frac{R^2}{H^2}r^2\cos^2\phi_1$$

o bien

$$\phi_1 = \arctan(R/H)$$

(exactamente igual que en el caso de $\mathbb{R}^3$). El plano $x_n = H$ se ve como

$$r\cos\phi_1 = H$$

o bien

$$r = \frac{H}{\cos\phi_1}$$

La región Ω' que corresponde al cono considerado es entonces

$$\begin{aligned}
&0 \le r \le \frac{H}{\cos\phi_1}, \quad 0 \le \phi_1 \le \arctan(R/H), \\
&0 \le \phi_i \le \pi, \quad i = 2, 3, \ldots, n-2, \quad 0 \le \phi_{n-1} \le 2\pi
\end{aligned}$$

y el volumen procurado, que denotaremos por $V(C^n)$, es

$$V(C^n) = \iint_{\Omega'} \cdots \int r^{n-1} \operatorname{sen}^{n-2}\phi_1 \operatorname{sen}^{n-3}\phi_2 \ldots \operatorname{sen}\phi_{n-2}\, dr\, d\phi_1 d\phi_2 \ldots d\phi_{n-1}$$

$$= \int_0^{\arctan(R/H)} \operatorname{sen}^{n-2}\phi_1\, d\phi_1 \int_0^{\pi} \operatorname{sen}^{n-3}\phi_2\, d\phi_2 \ldots \int_0^{\pi} \operatorname{sen}\phi_{n-2}\, d\phi_{n-2}$$

$$\int_0^{2\pi} d\phi_{n-1} \int_0^{H/\cos\phi_1} r^{n-1}\, dr$$

$$= \frac{H^n}{n}(2\pi) \int_0^{\arctan(R/H)} \frac{\operatorname{sen}^{n-2}\phi_1}{\cos^n\phi_1}\, d\phi_1 \int_0^{\pi} \operatorname{sen}^{n-3}\phi_2\, d\phi_2 \cdots \int_0^{\pi} \operatorname{sen}\phi_{n-2}\, d\phi_{n-2}$$

La primera de estas integrales se calcula fácilmente escribiendo el integrando como $\tan^{n-2}\phi_1 \sec^2\phi_1$. Su valor es $\frac{1}{n-1}\frac{R^{n-1}}{H^{n-1}}$. Por otra parte observe que el producto de las últimas $(n-3)$ integrales es igual a $\frac{n-1}{2\pi}V(S_{n-1})$ —ver ecuación (*) en el ejemplo anterior—, de modo que el volumen del cono en $\mathbb{R}^n$ $(n \geq 2)$ puede quedar escrito en términos del volumen de la esfera en $\mathbb{R}^{n-1}$ como

$$V(C^n) = \frac{H^n}{n}(2\pi)\frac{1}{n-1}\frac{R^{n-1}}{H^{n-1}}\frac{n-1}{2\pi}V(S_{n-1})$$

o sea

$$\boxed{V(C^n) = \frac{1}{n}R^{n-1}HV(s_{n-1})}$$

Por ejemplo, si $n = 2$ se tiene

$$V(C^2) = \frac{1}{2}RHV(S_1) = \frac{1}{2}RH(2) = RH$$

que es el área de un triángulo de base $2R$ y altura H. Si $n = 3$, se tiene

$$V(C^3) = \frac{1}{3}R^2HV(S_2) = \frac{1}{3}R^2H(\pi) = \frac{1}{3}\pi R^2 H$$

que es el volumen del cono en $\mathbb{R}^3$ que calculamos al comienzo del ejemplo. Si $n = 4$ nos queda que el volumen del cono en $\mathbb{R}^4$ es

$$V(C^4) = \frac{1}{4}R^3HV(S_3) = \frac{1}{4}R^3H\frac{4}{3}\pi = \frac{1}{3}\pi R^3 H$$

y el volumen del cono en $\mathbb{R}^5$ es

$$V(C^5) = \frac{1}{5}R^4HV(S_4) = \frac{1}{5}R^4H\frac{\pi^2}{2} = \frac{1}{10}\pi^2 R^4 H$$

etc. ∎

Ejercicios (Capítulo 6, Sección 9)

1. Sean $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ dos funciones continuas definidas en el conjunto abierto U. Sea Ω una región en $\mathbb{R}^n$ (como las consideradas en esta sección) contenida en U. Demuestre que

$$\iint_{\Omega}\cdots\int \left(f(x_1, x_2, \ldots, x_n) + cg(x_1, x_2, \ldots, x_n)\right) dx_1 dx_2 \ldots dx_n$$

$$= \iint_{\Omega}\cdots\int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n + c\iint_{\Omega}\cdots\int g(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$

donde c es un número real.

En los ejercicios 2—9, calcular las integrales múltiples indicadas.

2. $\iiiint_D x_1\, dx_1 dx_2 dx_3 dx_4$, donde $D = [0, 1] \times [1, 2] \times [2, 3] \times [3, 4]$.

3. $\iiiint_D x_1 x_2 x_3 x_4\, dx_1 dx_2 dx_3 dx_4$, donde $D = [0, 1]^4$.

4. $\iiiint_D (x_1 + x_2 + x_3 + x_4)\, dx_1 dx_2 dx_3 dx_4$, donde $D = [0, 1]^4$.

5. $\iiiint_D (x_1 + x_2 + x_3 + x_4)^2\, dx_1 dx_2 dx_3 dx_4$, donde D es la región limitada por los hiperplanos coordenados ($x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$), y el hiperplano $x_1 + x_2 + x_3 + x_4 = 1$.

6. $\int_0^1 \int_0^{x_1} \int_0^{x_2} \int_0^{x_3} dx_1 dx_2 dx_3 dx_4$

7. $\int_0^1 \int_0^{x_1} \int_0^{x_2} \int_0^{x_3} x_1 x_2 x_3 x_4\, dx_1 dx_2 dx_3 dx_4$

8. $\int_0^1 \int_0^{x_1} \int_0^{x_2} \int_0^{x_3} \int_0^{x_4} (x_1 + x_2 + x_3 + x_4 + x_5)\, dx_1 dx_2 dx_3 dx_4 dx_5$

9. $\int_0^{\pi} \int_0^{\pi} \int_0^{\pi} \int_0^{\pi} \int_0^{\pi} \int_0^{\pi} (x_1 + x_2)(x_3 + x_4)(x_5 + x_6)\, dx_1 dx_2 dx_3 dx_4 dx_5 dx_6$

10. Establezca una generalización del teorema del valor medio presentado para integrales dobles (sección 6.5.4) y para integrales triples (sección 6.8.3), en el caso de integrales n-múltiples.

11. Calcule el valor promedio del producto de cuatro números, si cada uno de estos varía en el intervalo [0, 1].

12. Calcule el valor promedio de la suma de los cuadrados de cinco números no negativos, si estos varían de tal modo que su suma nunca excede a la unidad.

13. (Un ejercicio sobre la función gamma). Se define la función gamma de variable x, denotada por $\Gamma(x)$, como la integral impropia

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1}\, dt$$

para valores de x cuya integral es convergente, es decir, para los cuales existe el límite

$$\lim_{b\to\infty} \int_0^{b} e^{-t} t^{x-1}\, dt$$

a. Demuestre que $\Gamma(1) = \Gamma(2) = 1$.

b. Demuestre que $\Gamma(1/2) = \sqrt{\pi}$. (Sugerencia: en la integral que define a $\Gamma(1/2)$ haga el cambio de variable $t = u^2$, para obtener que $\Gamma(1/2)$ es dos veces la integral $\int_0^{\infty} e^{-u^2}\, du$. Use ahora el resultado del ejemplo 8 de la sección 4).

c. Demuestre que para $x > 0$ se tiene que $\Gamma(x+1) = x\Gamma(x)$. (Sugerencia: en la integral que define a $\Gamma(x)$ haga integración por partes, escogiendo $u = e^{-t}$, $dv = t^{x-1}dt$).

d. Demuestre que para $n \in \mathbb{N}$, $\Gamma(n+1) = n!$. (Sugerencia: use el resultado del inciso anterior). Por esta propiedad a la función gamma se le suele llamar "función factorial generalizada".

e. Demuestre que

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(1)(3)(5)\cdots(2n-1)}{2^n}\sqrt{\pi}$$

donde $n \in \mathbb{N}$.

f. Demuestre que para $p > 0$, $\Gamma(x) = p^x \int_0^{\infty} e^{-ps} s^{x-1}\, ds$. (Sugerencia: en la integral que define a $\Gamma(x)$, haga $t = ps$, donde s es una nueva variable.)

14. (Un ejercicio sobre la función beta). Se define la función beta de las variables m y n, denotada por $B(m, n)$, como la integral impropia

$$B(m, n) = \int_0^{1} x^{m-1}(1-x)^{n-1}\, dx$$

para valores de m y n cuya integral converge (por ejemplo, —se puede demostrar que tal integral converge— para valores positivos de m y n).

a. Demuestre que $B(m, n) = B(n, m)$.

b. Demuestre que $B(m, n) = 2\int_0^{\pi/2} \operatorname{sen}^{2m-1}\theta \cos^{2n-1}\theta\, d\theta$. (Sugerencia: en la integral que define a $B(m, n)$ haga el cambio de variable $x = \operatorname{sen}^2\theta$).

c. Demuestre que $B(m, n) = \int_0^{\infty} \frac{u^{m-1}}{(1+u)^{m+n}}\, du$. (Sugerencia: en la integral que define a $B(m, n)$, haga el cambio de variable $u = \frac{x}{1-x}$).

d. Con el resultado del inciso **d** del ejercicio anterior demuestre que

$$\frac{1}{(1+u)^{m+n}} = \frac{1}{\Gamma(m+n)} \int_0^{\infty} e^{-(1+u)s} s^{m+n-1}\, ds$$

Obtenga entonces, del inciso anterior, que

$$B(m,n) = \int_0^\infty \frac{u^{m-1}}{\Gamma(m+n)} \int_0^\infty e^{-(1+u)s} s^{m+n-1}\, du\, ds$$
$$= \frac{1}{\Gamma(m+n)} \int_0^\infty e^{-s} s^{m+n-1}\, ds \int_0^\infty e^{-us} u^{m-1}\, du$$

e. Concluya del inciso anterior (y del resultado del inciso **d.** del problema anterior) que

$$B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

15. Demuestre que

$$\int_0^{\pi/2} \operatorname{sen}^k x\, dx = \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}+1\right)}$$

Concluya entonces que el volumen de la esfera S_n se puede escribir como

$$V(S_n) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$$

Capítulo 7

Integrales de línea

En este capítulo emprenderemos una nueva generalización del concepto de integral estudiado en el primer curso de cálculo. Esta generalización tendrá un sentido distinto al considerado en el capítulo anterior, en el que se estudiaron las integrales de funciones reales de n variables. Las regiones donde se efectuaba el proceso de integración eran subconjuntos de su dominio (subconjuntos de $\mathbb{R}^n$) limitados por rectas (o planos) y/o gráficas de funciones continuas de k variables (con $k < n$). Las regiones sobre las que se efectuarán las integrales que estudiaremos en este capítulo serán *curvas en el espacio,* ya estudiadas en el capítulo 5. También las funciones que se integrarán serán de naturaleza distinta a los integrandos del capítulo anterior.

Los problemas relacionados con el cálculo del trabajo realizado por una fuerza para llevar un punto de un lugar a otro, siguiendo una trayectoria determinada en el espacio, así como los problemas relacionados con el cálculo de la masa total y los momentos de un cuerpo "unidimensional" que recibe su forma por una curva en el espacio, cuya función conocida le da la densidad (lineal) del cuerpo en cada punto, son problemas que llevan al planteamiento de integrales de línea. En este capítulo se estudiarán estos dos tipos de problemas particulares como aplicación de la teoría de las integrales de línea que desarrollaremos. Por otra parte, también en este capítulo se estudiará uno de los teoremas célebres del cálculo en $\mathbb{R}^n$: el Teorema de Green. Con él haremos una demostración rigurosa del Teorema de Cambio de Variables en integrales múltiples, estudiado en el capítulo anterior.

7.1 Curvas en el espacio: Resumen de hechos importantes

En esta sección recordaremos algunos hechos sobre curvas estudiados en el capítulo 5 y que aplicaremos en el transcurso del presente capítulo. Advertimos que, en contraste a como se han tratado todos los temas en el libro, esta sección podrá parecer muy densa en contenido. La intención es presentar solamente un resumen que sirva de referencia para el desarrollo posterior del capítulo. Los temas que aquí aparecen ya se expusieron en forma regular ("con menos densidad") en el capítulo 5. Ante cualquier duda, acuda el lector a dicho capítulo para releer con más detalle los temas expuestos.

Una función $\boldsymbol{\lambda}: I \to \mathbb{R}^n$ continua, definida en el intervalo I de $\mathbb{R}$, se llama *camino* (o trayectoria) en $\mathbb{R}^n$. Siguiendo la tradición, a los caminos en este capítulo, que fungirán como regiones de integración en las que llamaremos "integrales de linea", los denotaremos por letras griegas ($\boldsymbol{\lambda}$ y sus vecinas). Siendo un camino $\boldsymbol{\lambda}: I \to \mathbb{R}^n$ una función que toma valores en $\mathbb{R}^n$, podemos escribir, para cada $t \in I$, su imagen $\boldsymbol{\lambda}(t)$ como $\boldsymbol{\lambda}(t) = (\lambda_1(t), \lambda_2(t), \ldots, \lambda_n(t))$ donde $\lambda_i: I \to \mathbb{R}$ son las *funciones coordenadas* de $\boldsymbol{\lambda}$, $i = 1, 2, \ldots, n$. Escribimos también al camino $\boldsymbol{\lambda}: I \to \mathbb{R}^n$ como $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Al

conjunto de imágenes $\boldsymbol{\lambda}(I) \subset \mathbb{R}^n$ se le llama *traza* del camino $\boldsymbol{\lambda}$ (lo llamaremos también "*curva* $\boldsymbol{\lambda}$"). Si $I = [a, b]$, diremos que $\boldsymbol{\lambda}(a) \in \mathbb{R}^n$ es el *punto inicial* del camino $\boldsymbol{\lambda}$, y que $\boldsymbol{\lambda}(b) \in \mathbb{R}^n$ es su *punto final*. Si $\boldsymbol{\lambda}(a) = \boldsymbol{\lambda}(b)$, diremos que $\boldsymbol{\lambda}$ es un camino *cerrado*. Si $\boldsymbol{\lambda}$ es una función inyectiva, diremos que $\boldsymbol{\lambda}$ es un camino *simple*. Geométricamente esto significa que la curva $\boldsymbol{\lambda}([a, b]) \subset \mathbb{R}^n$ no tiene autointersecciones. (Hecho curioso: un camino que sea cerrado no puede ser simple —¿por qué?—; sin embargo, existen los caminos "cerrados simples", pero éstos requieren una definición adicional). Si $\boldsymbol{\lambda}(a) = \boldsymbol{\lambda}(b)$ y la restricción de $\boldsymbol{\lambda}$ al intervalo $[a, b)$ es inyectiva, diremos que $\boldsymbol{\lambda}$ es un camino *cerrado simple*. Por ejemplo, si $\xi: [a, b] \to \mathbb{R}$ es una función real continua definida en el intervalo $[a, b]$, el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ dado por $\boldsymbol{\lambda}(t) = (t, \xi(t))$ es un camino simple, pero no cerrado (¿por qué?). El camino $\boldsymbol{\mu}: [-\pi, \pi] \to \mathbb{R}^2$, $\boldsymbol{\mu}(t) = (\operatorname{sen} t, \operatorname{sen} 2t)$ es cerrado, pero no es un camino cerrado simple, ya que la restricción de $\boldsymbol{\mu}$ al intervalo $[-\pi, \pi)$ no es inyectiva ($\boldsymbol{\mu}(-\pi) = \boldsymbol{\mu}(0) = (0, 0)$). Por último, el camino $\boldsymbol{\rho}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\rho}(t) = (\cos t, \operatorname{sen} t)$ es cerrado simple.

Consideremos el camino $\boldsymbol{\lambda}: I \to \mathbb{R}^n$. Éste se puede escribir como $\boldsymbol{\lambda}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$, en el que $x_j(t)$ son funciones reales definidas en el intervalo I (las funciones coordenadas de $\boldsymbol{\lambda}$). Decimos que el camino $\boldsymbol{\lambda}$ es diferenciable en t cuando existe la derivada $\boldsymbol{\lambda}'(t)$, la cual se define como el vector de $\mathbb{R}^n$, $\boldsymbol{\lambda}'(t) = (x_1'(t), x_2'(t), \ldots, x_n'(t))$. En otras palabras, $\boldsymbol{\lambda}$ será un camino diferenciable cuando (y sólo cuando) todas y cada una de sus funciones coordenadas sea diferenciable. En tal caso, al vector $\boldsymbol{\lambda}'(t)$ se le llama *vector velocidad* del camino $\boldsymbol{\lambda}$ en t, y es un vector "tangente" a la curva $\boldsymbol{\lambda}(I)$ en el punto $\boldsymbol{\lambda}(t)$. Se dice que el camino diferenciable $\boldsymbol{\lambda}: I \to \mathbb{R}^n$ es *regular* si $\boldsymbol{\lambda}'(t) \neq \mathbf{0}$ (¡el vector cero de $\mathbb{R}^n$!) $\forall t \in I$.

En el capítulo 5, donde se estudiaron de manera exhaustiva los caminos en el espacio, trabajamos sobre todo con caminos regulares. Esta propiedad de regularidad nos garantiza que el recorrido de las imágenes del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, que comienza en $\boldsymbol{\lambda}(a)$ y termina en $\boldsymbol{\lambda}(b)$, "no sufren detenciones —puntos en los que la velocidad $\boldsymbol{\lambda}'(t)$ es cero— en su recorrido"; es decir, comenzando en $\boldsymbol{\lambda}(a)$, las imágenes avanzan sobre la carretera $\boldsymbol{\lambda}([a, b])$ yendo siempre hacia $\boldsymbol{\lambda}(b)$ sin tener ninguna parada en la carretera, ya sea para seguir luego en el camino hacia $\boldsymbol{\lambda}(b)$, o para hacer un "pequeño regreso" sobre algún tramo de esta carretera. En este capítulo seguiremos trabajando con caminos regulares y por vía de sencillez, cuando en este capítulo se diga "camino", significará "camino regular". El principal interés estará en considerar caminos de clase $\mathscr{C}^1$. Esto significa que la función $\boldsymbol{\lambda}$ es diferenciable y tiene derivada $\boldsymbol{\lambda}'$ continua (lo que significa a su vez que cada una de sus funciones componentes $x_j'(t)$ es continua). Se considerarán también caminos por sección $\mathscr{C}^1$, es decir, caminos diferenciables $\boldsymbol{\lambda}: I \to \mathbb{R}^n$ ya que se da una partición P del intervalo I, digamos en k subintervalos $I_1, I_2, \ldots, I_k$ con la propiedad de que la función $\boldsymbol{\lambda}$ restringida a cada uno de estos subintervalos es de clase $\mathscr{C}^1$.

Recordemos el concepto de reparametrización de un camino. Sea $\boldsymbol{\lambda}: I \to \mathbb{R}^n$ un camino en $\mathbb{R}^n$ y $\phi: J \subseteq \mathbb{R} \to I$ una función de clase $\mathscr{C}^1$ sobreyectiva de modo que $\phi'(s) \neq 0$ $\forall s \in J$. Al camino compuesto $\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda} \circ \phi: J \subseteq \mathbb{R} \to \mathbb{R}^n$ se le llama *reparametrización* del camino $\boldsymbol{\lambda}$. Más aún, si $\phi'(s) > 0$ $\forall s \in J$, decimos que $\tilde{\boldsymbol{\lambda}}$ es una reparametrización de $\boldsymbol{\lambda}$ que *conserva la orientación*, mientras que si $\phi'(s) < 0$ $\forall s \in J$, se dice que $\tilde{\boldsymbol{\lambda}}$ es una reparametrización de $\boldsymbol{\lambda}$ que *invierte la orientación*. En el primer caso los caminos $\boldsymbol{\lambda}$ y $\tilde{\boldsymbol{\lambda}}$ tienen sus puntos inicial y final iguales (y recorren la traza —común a ambos— en $\mathbb{R}^n$ en la misma dirección), en tanto que en el segundo caso el punto inicial de $\boldsymbol{\lambda}$ es el punto final de $\tilde{\boldsymbol{\lambda}}$ y el punto final de $\boldsymbol{\lambda}$ es el punto inicial de $\tilde{\boldsymbol{\lambda}}$ (por lo que $\tilde{\boldsymbol{\lambda}}$ recorre la traza común a ambos en dirección contraria a $\boldsymbol{\lambda}$). En alguna parte de este capítulo consideraremos también composiciones $\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda} \circ \phi$ del camino $\boldsymbol{\lambda}$ con funciones $\phi: J \subseteq \mathbb{R} \to I$ de clase $\mathscr{C}^1$ sobreyectivas, las cuales no cumplen de modo necesario con la condición $\phi'(s) \neq 0$ $\forall s \in J$. Éstas no son, según la definición dada anteriormente, reparametrizaciones del camino $\boldsymbol{\lambda}$, pero se mencionarán cuando estudiemos una de las propiedades de las integrales de línea en la

siguiente sección. (Observe que siendo $\boldsymbol{\lambda}$ un camino regular, la composición de $\boldsymbol{\lambda}$ con una función ϕ de este tipo produce un nuevo camino que puede ya no ser regular, pues si por ejemplo $\phi'(s_0) = 0$, $s_0 \in J$, se tiene $\tilde{\boldsymbol{\lambda}}'(s_0) = (\boldsymbol{\lambda} \circ \phi)'(s_0) = \boldsymbol{\lambda}'(\phi(s_0))\phi'(s_0) = 0$).

Si $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ es un camino en $\mathbb{R}^n$, se puede componer con la función lineal sobreyectiva $\phi: [a, b] \to [a, b]$, $\phi(t) = a + b - t$ y obtener así una reparametrización de él, a saber $\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda} \circ \phi: [a, b] \to \mathbb{R}^n$, $\tilde{\boldsymbol{\lambda}}(t) = \boldsymbol{\lambda}(a + b - t)$. Obsérvese que: punto inicial de $\boldsymbol{\lambda} = \boldsymbol{\lambda}(a) = \tilde{\boldsymbol{\lambda}}(b) =$ punto final de $\tilde{\boldsymbol{\lambda}}$, y punto final de $\boldsymbol{\lambda} = \boldsymbol{\lambda}(b) = \tilde{\boldsymbol{\lambda}}(a) =$ punto inicial de $\tilde{\boldsymbol{\lambda}}$. De hecho, se puede ver que el camino $\tilde{\boldsymbol{\lambda}}$ recorre la misma curva que $\boldsymbol{\lambda}$ (pues $\boldsymbol{\lambda}([a, b]) = (-\boldsymbol{\lambda})([a, b])$) con la misma rapidez —magnitud de la velocidad— (pues $\|\tilde{\boldsymbol{\lambda}}'(t)\| = \|(\boldsymbol{\lambda}(a+b-t))'\| = \|\boldsymbol{\lambda}'(a+b-t)(-1)\| = \|\boldsymbol{\lambda}'(a+b-t)\|$) pero en sentido inverso. Escribiremos $\tilde{\boldsymbol{\lambda}} = -\boldsymbol{\lambda}$ y lo llamaremos "camino inverso de $\boldsymbol{\lambda}$".

Dado un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, diremos que éste es la *suma* de los dos caminos $\boldsymbol{\lambda}_1$ y $\boldsymbol{\lambda}_2$, lo cual se escribirá $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, si hay una $c \in [a, b]$ tal que la restricción de $\boldsymbol{\lambda}$ a $[a, c]$ coincida con $\boldsymbol{\lambda}_1$ y la restricción de $\boldsymbol{\lambda}$ a $[c, b]$ con $\boldsymbol{\lambda}_2$. En forma recíproca, con los dos caminos $\boldsymbol{\lambda}_1: [a, c] \to \mathbb{R}^n$ y $\boldsymbol{\lambda}_2: [c, b] \to \mathbb{R}^n$, para que $\boldsymbol{\lambda}_1(c) = \boldsymbol{\lambda}_2(c)$, podemos formar el camino suma $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2: [a, b] \to \mathbb{R}^n$ definido como

$$\boldsymbol{\lambda}(t) = \begin{cases} \boldsymbol{\lambda}_1(t) & \text{si } t \in [a, c] \\ \boldsymbol{\lambda}_2(t) & \text{si } t \in [c, b] \end{cases}$$

Más en general, con los dos caminos $\boldsymbol{\lambda}_1$ y $\boldsymbol{\lambda}_2$, podemos formar el camino suma $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$ con la única condición de que el punto final de $\boldsymbol{\lambda}_1$ coincida con el punto inicial de $\boldsymbol{\lambda}_2$. Es decir que si $\boldsymbol{\lambda}_1$ está definido en $[a, \alpha]$ y $\boldsymbol{\lambda}_2$ está definido en $[\beta, b]$, entonces (veremos que) podremos formar el camino suma $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$ con la única condición de que $\boldsymbol{\lambda}_1(\alpha) = \boldsymbol{\lambda}_2(\beta)$, sin ser necesariamente α igual a β. Para ello, reparametrizamos el camino $\boldsymbol{\lambda}_2$ componiéndolo con la función $\phi: [\alpha, \alpha + b - \beta] \to [\beta, b]$ dada por $\phi(s) = s - \alpha + \beta$. Sea $\tilde{\boldsymbol{\lambda}}_2: [\alpha, \alpha + b - \beta] \to \mathbb{R}^n$ tal reparametrización. Defina entonces $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$ como $\boldsymbol{\lambda}_1 + \tilde{\boldsymbol{\lambda}}_2: [a, \alpha + b - \beta] \to \mathbb{R}^n$.

En particular, obsérvese que para cualquier camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, tiene sentido hacer la suma $\boldsymbol{\lambda} + (-\boldsymbol{\lambda})$, donde $(-\boldsymbol{\lambda})$ es el camino inverso de $\boldsymbol{\lambda}$, y, más aún, que éste será un camino cerrado. Por otra parte, si $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ es un camino cerrado, podemos descomponerlo como $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$ escogiendo una $c \in [a, b]$ y definiendo $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}|_{[a,c]}$, $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}|_{[c,b]}$. Este hecho se aplicará en la demostración del teorema 7.4.1 en la sección 4. Nótese además que, en tal caso, $\boldsymbol{\lambda}_1(a) = -\boldsymbol{\lambda}_2(b)$ y $\boldsymbol{\lambda}_1(c) = -\boldsymbol{\lambda}_2(c)$, de modo que los caminos $\boldsymbol{\lambda}_1$ y $-\boldsymbol{\lambda}_2$ comparten su punto inicial y su punto final.

Si $\boldsymbol{\lambda}_1$ y $\boldsymbol{\lambda}_2$ son caminos de clase $\mathscr{C}^1$, con los que se puede formar el camino suma $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, este nuevo camino no es necesariamente también de clase $\mathscr{C}^1$. En el ejemplo 9 de la sección 5 del capítulo 5 se muestra un caso concreto que ejemplifica este hecho. Lo que sí resulta fácil de ver es que si $\boldsymbol{\lambda}_1$ y $\boldsymbol{\lambda}_2$ son caminos de clase $\mathscr{C}^1$ (o seccionalmente $\mathscr{C}^1$), el camino suma $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$ es seccionalmente $\mathscr{C}^1$. Más aún, el camino $\boldsymbol{\lambda}$ es seccionalmente $\mathscr{C}^1$ si y sólo si podemos escribir $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \cdots + \boldsymbol{\lambda}_k$ donde cada $\boldsymbol{\lambda}_i$ es un camino de clase $\mathscr{C}^1$.

7.2 Campos vectoriales

En este capítulo vamos a trabajar con funciones del tipo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, definidas en un abierto U de $\mathbb{R}^n$ y tomando valores en $\mathbb{R}^n$. Estas funciones serán los integrandos en las integrales que estudiaremos aquí. El objetivo de esta sección es familiarizarnos con estas funciones, tanto en su naturaleza y propiedades como visualizaciones geométricas, analizando algunos ejemplos importantes de este tipo de ellas. Comenzaremos por introducir la nomenclatura y recordar algunos hechos básicos sobre estas funciones.

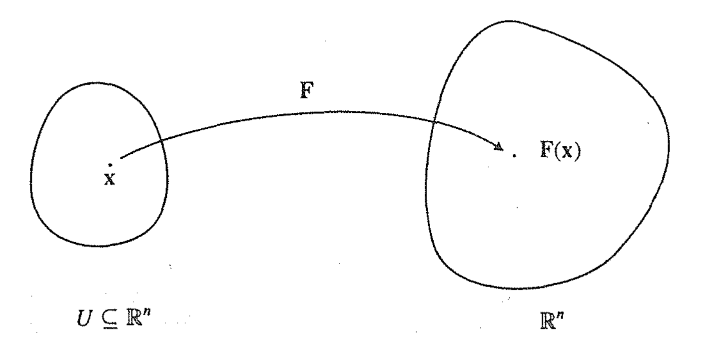


Figura 1. El campo vectorial $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$.

Una función del tipo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ se llama CAMPO VECTORIAL (en $\mathbb{R}^n$). Asocia a cada vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in U$, un vector $F(\mathbf{x}) \in \mathbb{R}^n$. De modo que podemos escribir la imagen $F(\mathbf{x}) \in \mathbb{R}^n$ del vector $\mathbf{x} \in U$ como $F(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_n(\mathbf{x}))$ donde cada $F_i(\mathbf{x})$ es un número real. Así pues, se tienen n funciones *reales* definidas en $U \subseteq \mathbb{R}^n$, $F_i: U \subseteq \mathbb{R}^n \to \mathbb{R}$, de modo que $F_i(\mathbf{x})$ es la i-ésima coordenada de la imagen $F(\mathbf{x}) \in \mathbb{R}^n$ del vector $\mathbf{x} \in U$. Se escribe entonces $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ y se dice que las funciones F_i son las *funciones coordenadas* del campo $\mathbf{F}$. En esquema se tiene

Una práctica muy popular para visualizar campos vectoriales es la siguiente: el campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ asocia a cada *punto* $\mathbf{x}$ *del espacio* $\mathbb{R}^n$ (más en concreto, a cada punto de $U \subseteq \mathbb{R}^n$), *el vector* $\mathbf{F}(\mathbf{x})$ *de* $\mathbb{R}^n$. (Observe que hemos jugado con la naturaleza "dual" de los elementos de $\mathbb{R}^n$: los elementos del dominio $\mathbf{x} \in U$ se ven como "puntos" en el espacio, y los del rango como "vectores" del espacio —ver sección 1 del capítulo 1). El vector $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$ se puede ver como una flecha cuyo su punto inicial está en el origen (el vector $\mathbf{0}$ de $\mathbb{R}^n$) y su punto final en el punto $\mathbf{F}(\mathbf{x})$. Más aún, esta flecha se puede colocar, mediante un movimiento rígido, donde queramos, y siempre representará al vector $\mathbf{F}(\mathbf{x})$. La idea para tener una visualización del campo F es entonces colocar la flecha $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$ de manera que su punto inicial sea $\mathbf{x} \in U$. Así, en una misma representación del espacio $\mathbb{R}^n$ veremos los puntos $\mathbf{x}$ del dominio del campo $\mathbf{F}$, y las imágenes de tales puntos como flechas que parten de ellos. Un esquema nos da

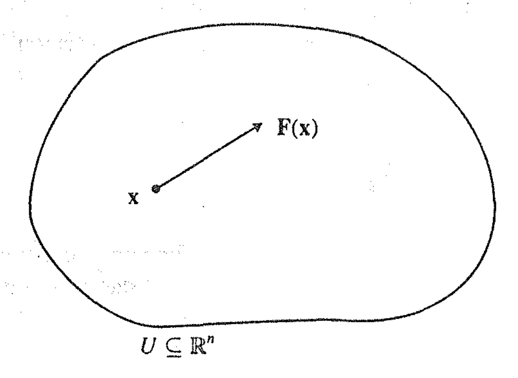


Figura 2. Representación del campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$.

Un campo en $\mathbb{R}^2$, $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ se verá entonces como un conjunto de flechas dibujadas dentro de U. Cada flecha es la imagen (que es un vector de $\mathbb{R}^2$) bajo el campo F del punto $\mathbf{x} \in U$ donde comienza la flecha. De modo semejante, un campo en $\mathbb{R}^3$, $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ se ve como un conjunto de flechas en el espacio interior de U.

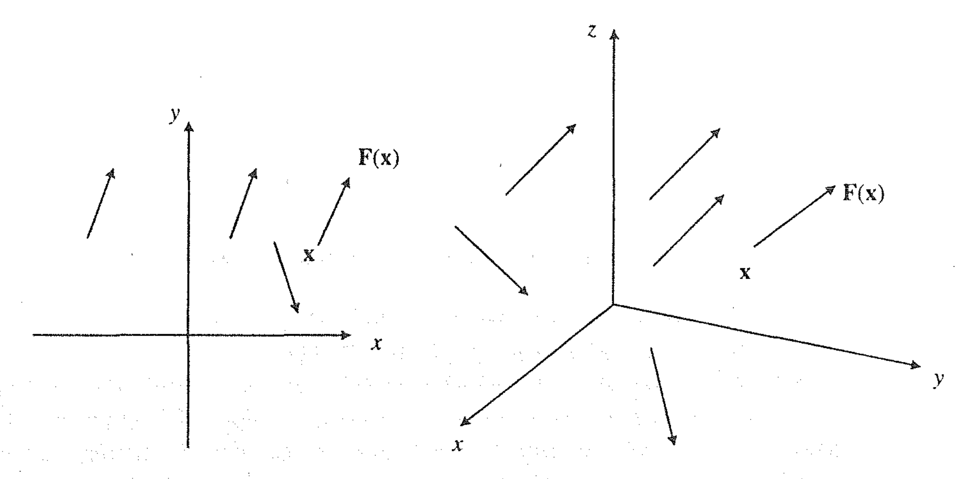


Figura 3. Campos vectoriales en $\mathbb{R}^2$ y en $\mathbb{R}^3$.

Esta idea para representar campos vectoriales es muy útil, pues se puede usar para describir, por ejemplo, *campos de velocidades*: piense en un líquido que fluya dentro de un tubo. En un momento dado, cada partícula del líquido se está moviendo a determinada velocidad (que es una magnitud vectorial) en la dirección del flujo. Así pues, podemos establecer la función $\mathbf{F}$: {puntos del interior del tubo} $\to \mathbb{R}^3$ que a cada partícula del fluido le asocie la velocidad que ésta tiene. Este es un campo en $\mathbb{R}^3$, llamado "campo de velocidades". Esquema

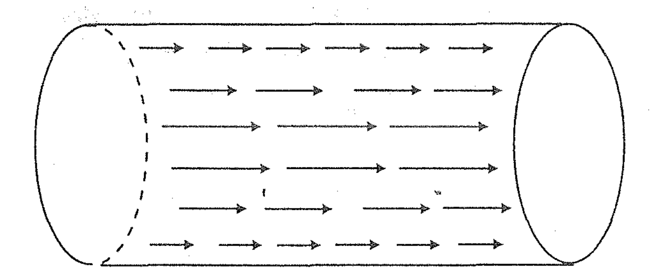


Figura 4. Campo de velocidades de un líquido que fluye por un tubo.

Del mismo modo, un *campo de fuerzas,* es un campo vectorial en $\mathbb{R}^2$ o en $\mathbb{R}^3$, donde en cada punto del plano ò del espacio se ejerce una determinada fuerza (que es una magnitud vectorial), de modo que se establece la función $\mathbf{F}: \mathbb{R}^2$ o $\mathbb{R}^3 \to \mathbb{R}^2$ o $\mathbb{R}^3$, en que a cada $\mathbf{x}$ le asocia la fuerza $\mathbf{F}(\mathbf{x})$ ejercida en $\mathbf{x}$. Un ejemplo de este tipo de campos es el campo gravitacional, en el que cada objeto de masa unitaria en el espacio circundante a la Tierra es atraído por una fuerza hacia ella, obedeciendo la Ley de Gravitación Universal. Dicha fuerza es proporcional al producto de la masa del objeto (que se supuso 1) y la masa de la Tierra, e inversamente proporcional al cuadrado de la distancia del

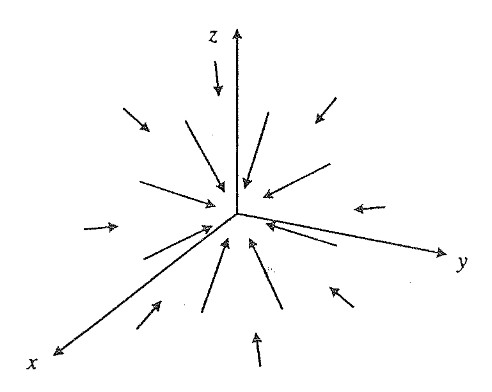


Figura 5. El campo gravitacional producido por la Tierra.

objeto dado y el centro de la Tierra (suponiendo que ahí es donde se encuentra el centro de masa de nuestro planeta). En un esquema se aprecia en la figura 5.

Otro ejemplo importante de campo vectorial, es el *campo electrostático* producido por una carga eléctrica q: supongamos que la carga q se ubica en el origen de coordenadas. En cada punto $\mathbf{x}$ del espacio $\mathbb{R}^3$ se coloca una carga de prueba positiva q_0. Esta será atraída o repelida por la carga q que obedece a la Ley de Coulomb, según la cual estas dos cargas experimentarán una fuerza $\mathbf{F}_\mathbf{x}$ (de atracción si q es negativa o de repulsión si q es positiva) directamente proporcional al producto de las cargas q y q_0 e inversamente proporcional al cuadrado de la distancia que las separa. El vector se define *campo eléctrico* en el punto $\mathbf{x}$, denotado por $\mathbf{E}(\mathbf{x})$, como la fuerza $\mathbf{F}_\mathbf{x}$ dividida por la carga q_o (que se había colocado en el punto $\mathbf{x}$). Nótese entonces que $\mathbf{E}(\mathbf{x})$ *depende sólo* (obviamente del valor de la carga q y) *del punto* $\mathbf{x}$ y *no de la carga* q_0 *colocada en ese punto*, pues

$$\mathbf{E}(\mathbf{x}) = \frac{\mathbf{F}_\mathbf{x}}{q_0} = \frac{\frac{kqq_0}{r^2}}{q_0} = \frac{kq}{r^2}$$

donde r es la distancia del punto $\mathbf{x}$ al origen de coordenadas (es decir, $r = \|\mathbf{x}\|$) y k es la constante de proporcionalidad de la ley de Coulomb. Podemos establecer entonces la función $\mathbf{E}: \mathbb{R}^3 \to \mathbb{R}^3$, de modo que a cada $\mathbf{x} \in \mathbb{R}^3$ se asocie el valor del campo eléctrico $\mathbf{E}(\mathbf{x})$ en el punto $\mathbf{x}$. Este es un campo vectorial en $\mathbb{R}^3$, cuyo aspecto gráfico se ve en la figura 6.

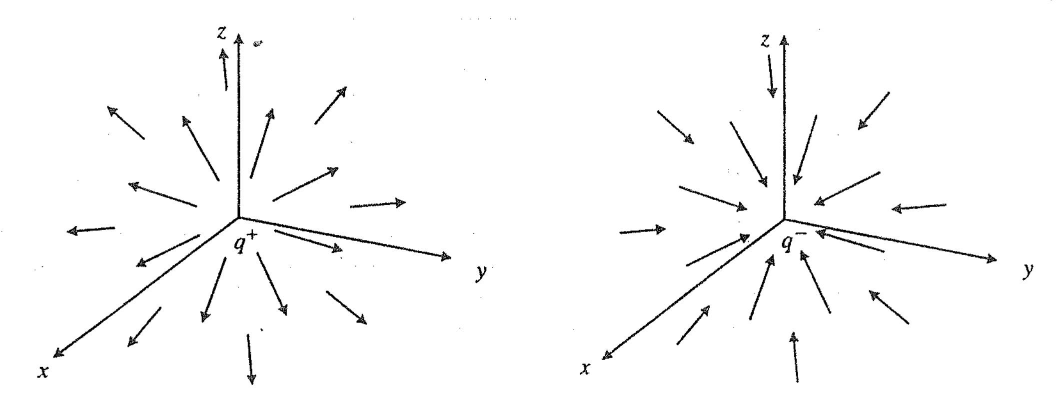


Figura 6. Campo eléctrico producido por una carga positiva y una negativa.

Se dice que el campo vectorial $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ es continuo (respectivamente diferenciable, o de clase $\mathscr{C}^k$) si (todas y cada una de) las funciones coordenadas $F_i: U \subseteq \mathbb{R}^n \to \mathbb{R}$ son continuas (diferenciables, o de clase $\mathscr{C}^k$). En este capítulo trabajaremos con campos vectoriales continuos (en su dominio).

Ejemplo 1. El campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (1, 0)$ es un campo constante que asocia a cada punto $(x, y) \in \mathbb{R}^2$ el mismo vector $(1, 0)$. Su aspecto gráfico es presenta en la figura 7.

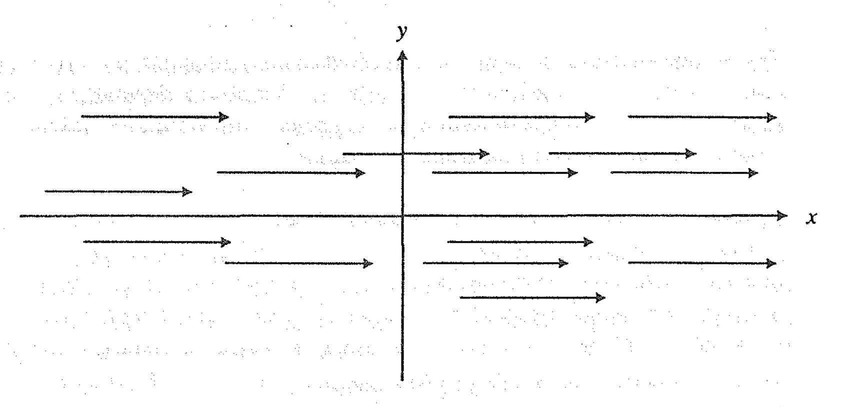


Figura 7. El campo constante $\mathbf{f}(x, y) = (1, 0)$ del ejemplo 1. ∎

Ejemplo 2. El campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x, y)$ que asocia al punto $(x, y) \in \mathbb{R}^2$ el vector $(x, y) \in \mathbb{R}^2$, se llama *campo radial* en $\mathbb{R}^2$. De manera análoga se tiene el campo radial en $\mathbb{R}^3$, $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $F(x, y, z) = (x, y, z)$, o, en general, en $\mathbb{R}^n$, $\mathbf{F}: \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F}(\mathbf{x}) = \mathbf{x}$. La imagen geométrica de este campo es un conjunto de flechas en dirección radial, como en la figura 8.

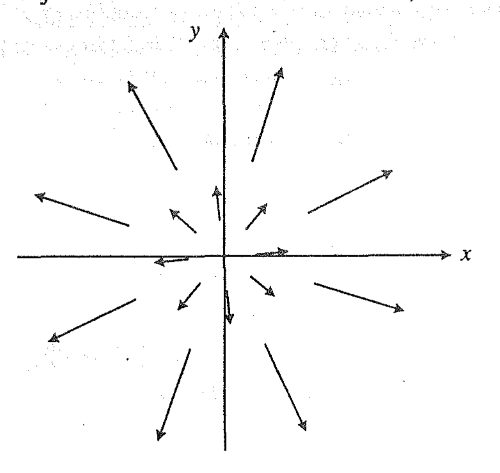


Figura 8. El campo radial $\mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F}(\mathbf{x}) = \mathbf{x}$. ∎

Para tener imágenes geométricas de campos en $\mathbb{R}^2$, en ocasiones es útil considerar las *líneas de campo*. Para un campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, una línea de campo se define como una curva en $\mathbb{R}^2$ cuya propiedad es que en cada punto de ella su tangente va en dirección del campo $\mathbf{F}$. Por ejemplo, para

el campo radial en $\mathbb{R}^2$, $\mathbf{F}(x, y) = (x, y)$, las líneas de campo son rectas que pasan por el origen de coordenadas. En general, para el campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (F_1(x, y), F_2(x, y))$, una línea de campo es una curva cuyo punto (x, y) tiene por derivada a la pendiente de la línea donde está $\mathbf{F}(x, y)$ (en el ejercicio 21 al final de esta sección se hace un estudio más preciso y detallado sobre estas ideas). Es fácil ver que esta pendiente es $\frac{F_2(x,y)}{F_1(x,y)}$, de modo que si $y = y(x)$ es la ecuación de la línea de campo, se debe tener

$$y' = \frac{F_2(x, y)}{F_1(x, y)}$$

Esta es una ecuación en la que se trata de despejar la función $y = y(x)$. A este tipo de ecuaciones, donde aparecen la "incógnita" $y = y(x)$, su variable independiente y su derivada, se les llama *ecuaciones diferenciales* (en este caso, ecuaciones diferenciales ordinarias de primer orden) y son objeto de estudios muy importantes en matemáticas.

Ejemplo 3. Para el campo del ejemplo 1, $\mathbf{F}(x, y) = (1, 0)$, en que $F_1(x, y) = 1$, $F_2(x, y) = 0$, las líneas de campo son líneas $y = y(x)$ que satisfacen la ecuación $y' = 0$. Integrando en ambos miembros esta ecuación obtenemos $y = c$. Es decir, las rectas paralelas al eje x son las líneas de campo del campo del ejemplo 1 (como se puede verificar fácilmente). Para el campo radial en $\mathbb{R}^2$, $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x, y)$, las líneas de campo satisfacen la ecuación $y' = \frac{y}{x}$. Aplicando $y' = \frac{dy}{dx}$, podemos escribir la ecuación anterior como $\frac{dy}{y} = \frac{dx}{x}$, de donde, por integración se obtiene $\ln y = \ln x + c$, y así c es una constante arbitraria. Ésta última expresión se puede escribir como $y = kx$, donde k es una constante. Es decir, las líneas del campo radial en $\mathbb{R}^2$ son rectas que pasan por el origen, como se puede ver fácilmente. ■

Ejemplo 4. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (-y, x)$. Las líneas de campo $y = y(x)$ deben satisfacer la ecuación $y' = -\frac{x}{y}$, o bien, $x\,dx + y\,dy = 0$. Integrando y multiplicando por 2 se obtiene $x^2 + y^2 = c$, donde c es una constante. Es decir, las líneas de campo son círculos concéntricos con centro en el origen de coordenadas. Otra observación que nos puede llevar a visualizar geométricamente este campo, es notar que el producto punto del vector (x, y) con el vector $F(x, y) = (-y, x)$ es cero, de modo que las flechas $\mathbf{F}(x, y)$ salen perpendiculares al vector (x, y) —el cual es un vector radial—, y por lo tanto, los vectores $\mathbf{F}(x, y)$ son tangentes a los círculos $x^2 + y^2 = c$, como ya se había concluido. La imagen geométrica de este campo se encuentra en la figura 9.

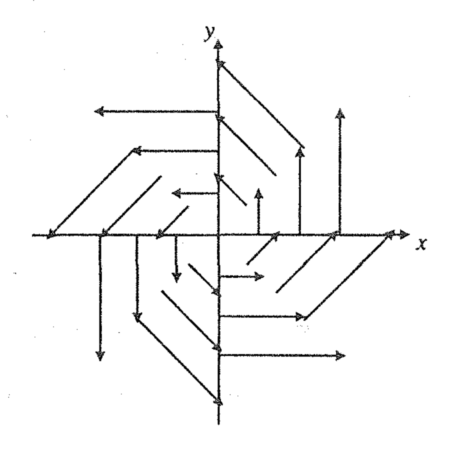


Figura 9. El campo $\mathbf{f}(x, y) = (-y, x)$. ■

Ejemplo 5. El campo $\mathbf{F}: \mathbb{R}^2 - \{(0,0)\} \to \mathbb{R}^2$, dado por

$$\mathbf{F}(x, y) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$$

tiene esencialmente el mismo aspecto del campo del ejemplo anterior, pues se conserva la propiedad de que $(x, y) \cdot \mathbf{F}(x, y) = 0$. La diferencia se encuentra sólo en el tamaño de las flechas que representan las imágenes $\mathbf{F}(x, y)$. En el campo del ejemplo anterior, la norma del vector $\mathbf{F}(x, y) = (-y, x)$ coincide con la norma del vector (x, y), de modo que a medida que nos alejamos del origen las flechas son de mayor tamaño, en tanto que en el campo de este ejemplo, la norma de las imágenes $\mathbf{F}(x, y)$ es $\|\mathbf{F}(x, y)\| = \frac{1}{\|(x,y)\|}$, de modo que en este caso las flechas serán tanto más pequeñas cuanto más nos alejemos del origen de coordenadas. El aspecto geométrico de este campo es entonces como se muestra en la figura 10.

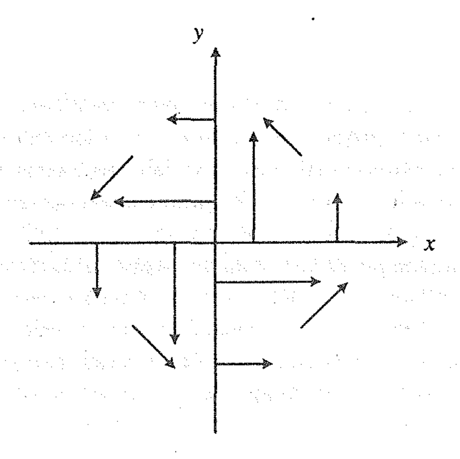


Figura 10. El campo del ejemplo 5.

Ejemplo 6. (Un ejemplo importante: el campo gradiente de una función de varias variables). Para comenzar este ejemplo, consideremos en primer lugar la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x$. Como sabemos, el gradiente de esta función en el punto (x, y) es el vector de $\mathbb{R}^2$ cuyas coordenadas son las derivadas parciales de la función. Éste es

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (1, 0)$$

Nótese que se puede establecer la función $\operatorname{grad} f: \mathbb{R}^2 \to \mathbb{R}^2$, que asocia a cada punto (x, y) el vector $\operatorname{grad} f(x, y)$. Éste es pues un campo vectorial en $\mathbb{R}^2$. En este caso, el campo es $\operatorname{grad} f(x, y) = (1, 0)$, que es el campo vectorial del ejemplo 1. En general, si consideramos una función real de n variables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable, podemos construir con ella el *campo vectorial gradiente* de f, $\operatorname{grad} f: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, que asocia a cada punto $\mathbf{x} \in U$, el vector $\operatorname{grad} f(\mathbf{x}) \in \mathbb{R}^n$. Por ejemplo, para la función $f: \mathbb{R}^n \to \mathbb{R}$, $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$, o bien $f(x_1, x_2, \ldots, x_n) = \frac{1}{2}(x_1^2 + x_2^2 + \cdots + x_n^2)$, tenemos que su campo gradiente es $\operatorname{grad} f: \mathbb{R}^n \to \mathbb{R}^n$,

$$\operatorname{grad} f(x_1, x_2, \ldots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) = (x_1, x_2, \ldots, x_n)$$

es decir, grad $f(\mathbf{x}) = \mathbf{x}$. Éste es el campo radial del ejemplo 2. De modo similar, para la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \operatorname{sen}(x^2 + y^2) + 3y^4$, su campo gradiente es

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2x \operatorname{sen}(x^2 + y^2), 2y \operatorname{sen}(x^2 + y^2) + 12y^3) \qquad \blacksquare$$

En base a las ideas presentadas en el ejemplo anterior, nos podríamos plantear la siguiente pregunta: dado un campo vectorial $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, ¿éste es el campo gradiente de alguna función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$? Es decir, ¿hay alguna función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ de modo que grad $f(\mathbf{x}) = \mathbf{F}(\mathbf{x})$, $\mathbf{x} \in U$? O de modo más explícito: ¿hay alguna función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ que

$$\frac{\partial f}{\partial x_1} = F_1, \quad \frac{\partial f}{\partial x_2} = F_2, \quad \ldots \quad , \frac{\partial f}{\partial x_n} = F_n$$

donde $F_1, F_2, \ldots, F_n$ son funciones reales definidas en $U \subseteq \mathbb{R}^n$ dadas? Llamamos la atención al contenido de esta pregunta en el caso $n = 1$: dado un campo vectorial en $\mathbb{R}$, $F: I \subseteq \mathbb{R} \to \mathbb{R}$ (ésta no es más que una función real de una variable real, de las que ya hace mucho tiempo conocemos), ¿hay alguna una función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ diferenciable de modo que grad $f(\mathbf{x}) = f'(\mathbf{x}) = F(\mathbf{x})$, $\mathbf{x} \in I$? La respuesta a esta pregunta la podría proporcionar cualquier estudiante de un primer curso de cálculo: la función buscada puede ser cualquier primitiva de $F(\mathbf{x})$, o bien, siendo más explícitos, con la ayuda del Teorema Fundamental del Cálculo, se puede escribir la respuesta como $f(x) = \int_a^x F(x)\,dx$, donde $a \in I$. Para $n > 1$, la pregunta se torna (como era de esperarse) de respuesta mucho más complicada que en el caso $n = 1$. En realidad, uno de los objetivos centrales de este capítulo es discutir la respuesta a esta pregunta, así como ver la importancia que ella tiene tanto dentro de la teoría, como de las aplicaciones que se estudiarán a lo largo del capítulo.

Apéndice Campos vectoriales en los sistemas de coordenadas cilíndricas y esféricas

Dentro de las aplicaciones físicas de la teoría que se desarrollará en este capítulo (por ejemplo en el estudio de campos eléctricos) es común trabajar con campos vectoriales en $\mathbb{R}^3$ descritos en otros sistemas coordenados distintos del cartesiano. En este apéndice vamos a obtener expresiones que nos den la correspondencia entre un campo $\mathbf{F}$ en $\mathbb{R}^3$, escrito en el sistema cartesiano, y su equivalente, escrito en los sistemas de coordenadas cilíndricas y esféricas.

Escribamos las funciones coordenadas del campo $\mathbf{F}$ como F_x, F_y y F_z. Entonces, el campo $\mathbf{F}$ en el sistema cartesiano se ve como

$$\begin{aligned}\mathbf{F}(x, y, z) &= (F_x(x, y, z), F_y(x, y, z), F_z(x, y, z)) \\ &= F_x(x, y, z)\mathbf{i} + F_y(x, y, z)\mathbf{j} + F_z(x, y, z)\mathbf{k}\end{aligned}$$

donde $\mathbf{i}, \mathbf{j}, \mathbf{k}$ son los vectores de la base canónica de $\mathbb{R}^3$. Se entiende así que F_x es una función real de las variables x, y, z que nos da la proyección de F en la dirección del vector $\mathbf{i} = (1, 0, 0)$ (análogamente para F_y y F_z). En el sistema de coordenadas cilíndricas, el espacio $\mathbb{R}^3$ se describe en

términos de la base ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z\}$, donde

$$\begin{aligned} \mathbf{e}_r &= \cos\theta\mathbf{i} + \operatorname{sen}\theta\mathbf{j} \\ \mathbf{e}_\theta &= -\operatorname{sen}\theta\mathbf{i} + \cos\theta\mathbf{j} \\ \mathbf{e}_z &= \mathbf{k} \end{aligned}$$

(ver apéndice de la sección 5 del capítulo 1). De aquí podemos obtener que

$$\begin{aligned} \mathbf{i} &= \cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta \\ \mathbf{j} &= \operatorname{sen}\theta\mathbf{e}_r + \cos\theta\mathbf{e}_\theta \\ \mathbf{k} &= \mathbf{e}_z \end{aligned}$$

En esquema estos vectores se ven como

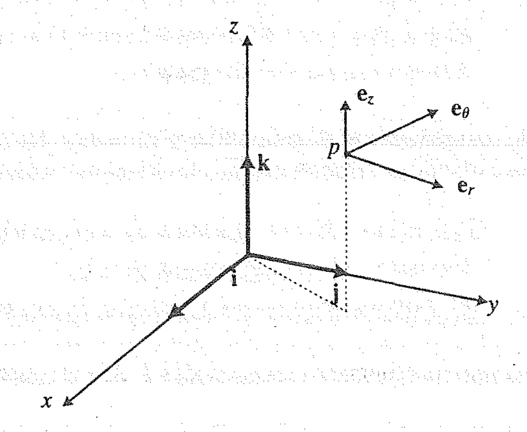


Figura 11. Vectores de las bases canónicas de $\mathbb{R}^3$ en los sistemas de coordenadas cartesianas y cilíndricas.

Las ecuaciones de transformación del sistema cilíndrico al cartesiano son, como sabemos, $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, $z = z$. Para las funciones coordenadas del campo $\mathbf{F}$, digamos F_x, la cual depende de x, y, z, escribimos, al pasar a las coordenadas cilíndricas, $\tilde{F}_x(r, \theta, z) = F_x(r\cos\theta, r\operatorname{sen}\theta, z)$. (la tilde es para denotar a la nueva función compuesta de F_x con las funciones $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, $z = z$). Es decir, el campo F se vería como

$$\begin{aligned} \tilde{\mathbf{F}}(r, \theta, z) &= (\tilde{F}_x(r, \theta, z), \tilde{F}_y(r, \theta, z), \tilde{F}_z(r, \theta, z)) \\ &= (F_x(r\cos\theta, r\operatorname{sen}\theta, z), F_y(r\cos\theta, r\operatorname{sen}\theta, z), F_z(r\cos\theta, r\operatorname{sen}\theta, z)) \end{aligned}$$

Ésta aún no es la expresión del campo $\mathbf{F}$ en coordenadas cilíndricas. Lo que hemos hecho hasta el momento es expresar los vectores (x, y, z) *del dominio* del campo $\mathbf{F}$ en términos de las coordenadas r, θ, z. Debemos ahora transformar los vectores *de las imágenes* del campo para escribirlos en

términos de la base $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_z$. Se tiene

$$\begin{aligned}\tilde{\mathbf{F}}(r,\theta,z) &= (\tilde{F}_x(r,\theta,z), \tilde{F}_y(r,\theta,z), \tilde{F}_z(r,\theta,z))\\ &= \tilde{F}_x(r,\theta,z)\mathbf{i} + \tilde{F}_y(r,\theta,z)\mathbf{j} + \tilde{F}_z(r,\theta,z))\mathbf{k}\\ &= \tilde{F}_x(r,\theta,z)(\cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta) + \tilde{F}_y(r,\theta,z)(\operatorname{sen}\theta\mathbf{e}_r + \cos\theta\mathbf{e}_\theta) + \tilde{F}_z(r,\theta,z)\mathbf{e}_z\\ &= (\cos\theta\tilde{F}_x(r,\theta,z) + \operatorname{sen}\theta\tilde{F}_y(r,\theta,z))\mathbf{e}_r\\ &\quad + (-\operatorname{sen}\theta\tilde{F}_x(r,\theta,z) + \cos\theta\tilde{F}_y(r,\theta,z))\mathbf{e}_\theta + \tilde{F}_z(r,\theta,z)\mathbf{e}_z\end{aligned}$$

Entonces, escribiendo las nuevas coordenadas del campo $\mathbf{F}$ en coordenadas cilíndricas como F_r, F_θ, F_z, es decir,

$$\begin{aligned}\tilde{\mathbf{F}}(r,\theta,z) &= (F_r(r,\theta,z), F_\theta(r,\theta,z), F_z(r,\theta,z))\\ &= F_r(r,\theta,z)\mathbf{e}_r + F_\theta(r,\theta,z)\mathbf{e}_\theta + F_z(r,\theta,z)\mathbf{e}_z\end{aligned}$$

tenemos que

$$\begin{aligned}F_r(r,\theta,z) &= \cos\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z) + \operatorname{sen}\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z)\\ F_\theta(r,\theta,z) &= -\operatorname{sen}\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z) + \cos\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z)\\ F_z(r,\theta,z) &= F_z(r\cos\theta, r\operatorname{sen}\theta, z)\end{aligned}$$

Por ejemplo, consideremos el campo $\mathbf{F}(x,y,z) = (xy, 2z, zx)$. Veamos cómo se ve este campo en coordenadas cilíndricas. Transformemos las funciones coordenadas de $\mathbf{F}$

$$\begin{aligned}\tilde{\mathbf{F}}_x(r,\theta,z) &= F_x(r\cos\theta, r\operatorname{sen}\theta, z) = (r\cos\theta)(r\operatorname{sen}\theta) = r^2\operatorname{sen}\theta\cos\theta\\ \tilde{F}_y(r,\theta,z) &= F_y(r\cos\theta, r\operatorname{sen}\theta, z) = 2z\\ \tilde{F}_z(r,\theta,z) &= F_z(r\cos\theta, r\operatorname{sen}\theta, z) = z(r\cos\theta) = rz\cos\theta\end{aligned}$$

Entonces las nuevas funciones coordenadas F_r, F_θ y F_z serán

$$\begin{aligned}F_r(r,\theta,z) &= (\cos\theta)(r^2\operatorname{sen}\theta\cos\theta) + (\operatorname{sen}\theta)(2z) = r^2\operatorname{sen}\theta\cos^2\theta + 2z\operatorname{sen}\theta\\ F_\theta(r,\theta,z) &= (-\operatorname{sen}\theta)(r^2\operatorname{sen}\theta\cos\theta) + (\cos\theta)(2z) = -r^2\operatorname{sen}^2\theta\cos\theta + 2z\cos\theta\\ F_z(r,\theta,z) &= rz\cos\theta\end{aligned}$$

y así, la expresión del campo $\mathbf{F}$ en coordenadas cilíndricas es

$$\begin{aligned}\tilde{F}(r,\theta,z) &= (r^2\operatorname{sen}\theta\cos^2\theta + 2z\operatorname{sen}\theta, -r^2\operatorname{sen}^2\theta\cos\theta + 2z\cos\theta, rz\cos\theta)\\ &= (r^2\operatorname{sen}\theta\cos^2\theta + 2z\operatorname{sen}\theta)\mathbf{e}_r + (-r^2\operatorname{sen}^2\theta\cos\theta + 2z\cos\theta)\mathbf{e}_\theta + (rz\cos\theta)\mathbf{e}_z\end{aligned}$$

El tratamiento que se da en el caso de las coordenadas esféricas es completamente similar. En este caso tenemos también una base ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ en la que los vectores marcan las direcciones de medición de las coordenadas esféricas r, θ, ϕ de un punto $\mathbf{p}$ en $\mathbb{R}^3$, como se muestra en la figura 12.

Necesitamos las fórmulas que relacionan los vectores $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ del sistema ortonormal de la base canónica de $\mathbb{R}^3$ con los vectores $\mathbf{e}_r$, $\mathbf{e}_\phi$, $\mathbf{e}_\theta$ del sistema ortonormal en coordenadas esféricas. Estas fórmulas se establecieron en el apéndice de la sección 5 del capítulo 1, pero, en atención al lector, recordamos brevemente las ideas generales de su deducción. Recuerde que r es la distancia de $\mathbf{p}$ al

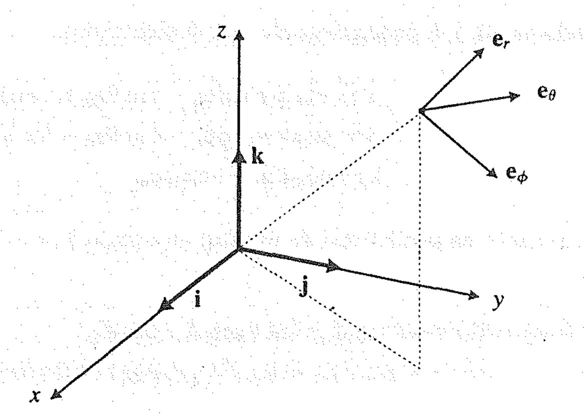


Figura 12. Los vectores de las bases canonícas de $\mathbb{R}^3$ en los sistemas de coordenadas cartesiano y esférico.

origen. Entonces el vector $\mathbf{e}_r$ debe ser un vector "radial unitario", es decir, si $\mathbf{p} = (x, y, z)$ se debe tener

$$\mathbf{e}_r = \frac{\mathbf{p}}{\|\mathbf{p}\|} = \frac{1}{\sqrt{x^2 + y^2 + z^2}}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$$

o bien, como las ecuaciones de transformación en el caso de las coordenadas esféricas son

$$x = r\cos\theta\,\mathrm{sen}\,\phi, \quad y = r\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\phi, \quad z = r\cos\phi$$

nos queda que

$$\mathbf{e}_r = \cos\theta\,\mathrm{sen}\,\phi\mathbf{i} + \mathrm{sen}\,\theta\,\mathrm{sen}\,\phi\mathbf{j} + \cos\phi\mathbf{k}$$

Recuerde también que la coordenada θ de las coordenadas esféricas es la misma coordenada θ del sistema cilíndrico, por lo que el vector $\mathbf{e}_\theta$ es el mismo en los dos sistemas coordenados (¿por qué?). Entonces

$$\mathbf{e}_\theta = -\,\mathrm{sen}\,\theta\mathbf{i} + \cos\theta\mathbf{j}$$

Por último, el vector restante e_ϕ lo podemos obtener usando el hecho de que éste debe ser ortogonal tanto a $\mathbf{e}_r$ como a $\mathbf{e}_\theta$. Haciendo entonces el producto cruz $\mathbf{e}_\theta \times \mathbf{e}_r$ obtenemos un vector *unitario* (¿por qué?) ortogonal a $\mathbf{e}_r$ y a $\mathbf{e}_\theta$, en la dirección "que baja" del eje z, como debe ser (pues así se mide la coordenada ϕ). Entonces

$$\mathbf{e}_\phi = \mathbf{e}_\theta \times \mathbf{e}_r = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\,\mathrm{sen}\,\theta & \cos\theta & 0 \\ \cos\theta\,\mathrm{sen}\,\phi & \mathrm{sen}\,\theta\,\mathrm{sen}\,\phi & \cos\phi \end{bmatrix}$$
$$= \cos\theta\cos\phi\mathbf{i} + \mathrm{sen}\,\theta\cos\phi\mathbf{j} - \mathrm{sen}\,\phi\mathbf{k}$$

En resumen, tenemos las expresiones de los vectores de la base ortonormal $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ en términos de los vectores de la base canónica $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$

$$\mathbf{e}_r = \cos\theta\,\mathrm{sen}\,\phi\mathbf{i} + \mathrm{sen}\,\theta\,\mathrm{sen}\,\phi\mathbf{j} + \cos\phi\mathbf{k}$$
$$\mathbf{e}_\theta = -\,\mathrm{sen}\,\theta\mathbf{i} + \cos\theta\mathbf{j}$$
$$\mathbf{e}_\phi = \cos\theta\cos\phi\mathbf{i} + \mathrm{sen}\,\theta\cos\phi\mathbf{j} - \mathrm{sen}\,\phi\mathbf{k}$$

Resolviendo para **i**, **j**, **k** en términos de r, θ, ϕ obtenemos

$$\begin{aligned}\mathbf{i} &= \operatorname{sen}\phi\cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta + \cos\phi\cos\theta\mathbf{e}_\phi\\ \mathbf{j} &= \operatorname{sen}\theta\operatorname{sen}\phi\mathbf{e}_r + \cos\theta\mathbf{e}_\theta + \operatorname{sen}\theta\cos\phi\mathbf{e}_\phi\\ \mathbf{k} &= \cos\phi\mathbf{e}_r - \operatorname{sen}\phi\mathbf{e}_\phi\end{aligned}$$

Ahora ya estamos en posibilidad de escribir un campo $\mathbf{F} = (F_x, F_y, F_z)$ en coordenadas esféricas. Tenemos

$$\begin{aligned}\tilde{\mathbf{F}}(r, \theta, \phi) &= \mathbf{F}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\\ &= (\tilde{F}_x(r,\theta,\phi), \tilde{F}_y(r,\theta,\phi), \tilde{F}_z(r,\theta,\phi)) = \tilde{F}_x(r,\theta,\phi)\mathbf{i} + \tilde{F}_y(r,\theta,\phi)\mathbf{j} + \tilde{F}_z(r,\theta,\phi)\mathbf{k}\end{aligned}$$

donde

$$\begin{aligned}\tilde{F}_x(r,\theta,\phi) &= F_x(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\\ \tilde{F}_y(r,\theta,\phi) &= F_y(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\\ \tilde{F}_z(r,\theta,\phi) &= F_z(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\end{aligned}$$

Sustituyendo las expresiones de **i**, **j** y **k** en términos de $\mathbf{e}_r$, $\mathbf{e}_\theta$ y $\mathbf{e}_\phi$, nos queda

$$\begin{aligned}\tilde{\mathbf{F}}(r,\theta,\phi) &= \tilde{F}_x(r,\theta,\phi)(\operatorname{sen}\phi\cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta + \cos\phi\cos\theta\mathbf{e}_\phi)\\ &\quad + \tilde{F}_y(r,\theta,\phi)(\operatorname{sen}\theta\operatorname{sen}\phi\mathbf{e}_r + \cos\theta\mathbf{e}_\theta + \operatorname{sen}\theta\cos\phi\mathbf{e}_\phi)\\ &\quad + \tilde{F}_z(r,\theta,\phi)(\cos\phi\mathbf{e}_r - \operatorname{sen}\phi\mathbf{e}_\phi)\\ &= (\operatorname{sen}\phi\cos\theta\tilde{F}_x(r,\theta,\phi) + \operatorname{sen}\theta\operatorname{sen}\phi\tilde{F}_y(r,\theta,\phi) + \cos\phi\tilde{F}_z(r,\theta,\phi))\mathbf{e}_r\\ &\quad + (-\operatorname{sen}\theta\tilde{F}_x(r,\theta,\phi) + \cos\theta\tilde{F}_y(r,\theta,\phi))\mathbf{e}_\theta\\ &\quad + (\cos\phi\cos\theta\tilde{F}_x(r,\theta,\phi) + \operatorname{sen}\theta\cos\phi\tilde{F}_y(r,\theta,\phi) - \operatorname{sen}\phi\tilde{F}_z(r,\theta,\phi))\mathbf{e}_\phi\end{aligned}$$

A manera de ejemplo, tomemos el campo

$$\mathbf{F}(x, y, z) = \left(x^2 + y^2 + z^2, \sqrt{x^2 + y^2}, z^2\right)$$

y escribámoslo en coordenadas esféricas. Las funciones coordenadas se ven como

$$\begin{aligned}\tilde{F}_x(r,\theta,\phi) &= F_x(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\\ &= (r\cos\theta\operatorname{sen}\phi)^2 + (r\operatorname{sen}\theta\operatorname{sen}\phi)^2 + (r\cos\phi)^2 = r^2\\ \tilde{F}_y(r,\theta,\phi) &= F_y(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\\ &= \sqrt{(r\cos\theta\operatorname{sen}\phi)^2 + (r\operatorname{sen}\theta\operatorname{sen}\phi)^2} = r\operatorname{sen}\phi\\ \tilde{F}_z(r,\theta,\phi) &= F_z(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi) = r\cos\phi\end{aligned}$$

Las nuevas funciones coordenadas son

$$\begin{aligned}F_r(r,\theta,\phi) &= \operatorname{sen}\phi\cos\theta\tilde{F}_x(r,\theta,\phi) + \operatorname{sen}\theta\operatorname{sen}\phi\tilde{F}_y(r,\theta,\phi) + \cos\phi\tilde{F}_z(r,\theta,\phi)\\ &= \operatorname{sen}\phi\cos\theta r^2 + \operatorname{sen}\theta\operatorname{sen}\phi r\operatorname{sen}\phi + \cos\phi r\cos\phi\\ &= r^2\operatorname{sen}\phi\cos\theta + r\operatorname{sen}\theta\operatorname{sen}^2\phi + r\cos^2\phi\end{aligned}$$

$$\begin{aligned}F_\theta(r,\theta,\phi) &= -\operatorname{sen}\theta\tilde{F}_x(r,\theta,\phi)+\cos\theta\tilde{F}_y(r,\theta,\phi)\\ &= -\operatorname{sen}\theta r^2+\cos\theta r\operatorname{sen}\phi = -r^2\operatorname{sen}\theta+r\cos\theta\operatorname{sen}\phi\\ F_\phi(r,\theta,\phi) &= \cos\phi\cos\theta\tilde{F}_x(r,\theta,\phi)+\operatorname{sen}\theta\cos\phi\tilde{F}_y(r,\theta,\phi)-\operatorname{sen}\phi\tilde{F}_z(r,\theta,\phi)\\ &= \cos\phi\cos\theta r^2+\operatorname{sen}\theta\cos\phi r\operatorname{sen}\phi-\operatorname{sen}\phi r\cos\phi\\ &= r^2\cos\theta\cos\phi+r\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi-r\operatorname{sen}\phi\cos\phi\end{aligned}$$

Entonces el campo **F** se ve en el sistema esférico como

$$\begin{aligned}\tilde{\mathbf{F}}(r,\theta,\phi) = {}&(r^2\cos\theta\operatorname{sen}\phi+r\operatorname{sen}\theta\operatorname{sen}^2\phi+r\cos^2\phi)\mathbf{e}_r\\ &+(-r^2\operatorname{sen}\theta+r\cos\theta\operatorname{sen}\phi)\mathbf{e}_\theta\\ &+(r^2\cos\theta\cos\phi+r\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi-r\operatorname{sen}\phi\cos\phi)\mathbf{e}_\phi\end{aligned}$$

A manera de resumen presentamos los resultados anteriores en la siguiente

Proposición Sea $\mathbf{F}=(F_x,F_y,F_z)$ un campo en $\mathbb{R}^3$.

a. La expresión del campo F en coordenadas cilídricas es

$$\begin{aligned}\mathbf{F}(r\cos\theta, r\operatorname{sen}\theta, z) = \tilde{\mathbf{F}}(r,\theta,z) &= (F_r(r,\theta,z), F_\theta(r,\theta,z), F_z(r,\theta.z))\\ &= F_r(r,\theta,z)\mathbf{e}_r+F_\theta(r,\theta,z)\mathbf{e}_\theta+F_z(r,\theta.z)\mathbf{e}_z\end{aligned}$$

donde

$$\begin{aligned}F_r(r,\theta,z) &= \cos\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z)+\operatorname{sen}\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z)\\ F_\theta(r,\theta,z) &= -\operatorname{sen}\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z)+\cos\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z)\\ F_z(r,\theta,z) &= F_z(r\cos\theta, r\operatorname{sen}\theta, z)\end{aligned}$$

b. La expresión del campo **F** en coordenadas esféricas es

$$\begin{aligned}\mathbf{F}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi) &= \tilde{\mathbf{F}}(r,\theta,\phi)\\ &= (F_r(r,\theta,\phi), F_\theta(r,\theta,\phi), F_\phi(r,\theta,\phi))\\ &= F_r(r,\theta,\phi)\mathbf{e}_r+F_\theta(r,\theta,\phi)\mathbf{e}_\theta+F_\phi(r,\theta,\phi)\mathbf{e}_\phi\end{aligned}$$

donde

$$\begin{aligned}F_r(r,\theta,\phi) &= \cos\theta\operatorname{sen}\phi\tilde{F}_x(r,\theta,\phi)+\operatorname{sen}\theta\operatorname{sen}\phi\tilde{F}_y(r,\theta,\phi)+\cos\phi\tilde{F}_z(r,\theta,\phi)\\ F_\theta(r,\theta,\phi) &= -\operatorname{sen}\theta\tilde{F}_x(r,\theta,\phi)+\cos\theta\tilde{F}_y(r,\theta,\phi)\\ F_\phi(r,\theta,\phi) &= \cos\phi\cos\theta\tilde{F}_x(r,\theta,\phi)+\operatorname{sen}\theta\cos\phi\tilde{F}_y(r,\theta,\phi)-\operatorname{sen}\phi\tilde{F}_z(r,\theta,\phi)\end{aligned}$$

donde a su vez

$$\tilde{F}_{x(y,z)}(r,\theta,\phi) = F_{x(y,z)}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$$

■

Ejercicios (Capítulo 7, Sección 2)

En cada uno de los ejercicios 1–8, bosqueje algunos de los vectores (imágenes) del campo $F: \mathbb{R}^2 \to \mathbb{R}^2$ dado.

1. $\mathbf{F}(x, y) = (0, 2)$

2. $\mathbf{F}(x, y) = (1, x)$

3. $\mathbf{F}(x, y) = (x, 1)$

4. $\mathbf{F}(x, y) = (x, x)$

5. $\mathbf{F}(x, y) = (y, y)$

6. $\mathbf{F}(x, y) = (x, 2y)$

7. $\mathbf{F}(x, y) = (-2y, x)$

8. $\mathbf{F}(x, y) = (-y, -x)$

9. Describa las líneas de campo para los campos de los ejercicios 4, 6 y 7.

En los ejercicios 10–14, obtenga el campo gradiente de la función diferenciable $f: \mathbb{R}^2 \to \mathbb{R}$ dada, y bosqueje algunos de sus elementos en el plano.

10. $f(x, y) = x + y$

11. $f(x, y) = xy$

12. $f(x, y) = x^2 + y^2$

13. $f(x, y) = 2x^2 - y^2$

14. $f(x, y) = x^2 - 2y^2$

15. Considere el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (-y, x, 0)$. Compruebe que en cada punto $P = (x, y, z)$ distinto del origen de coordenadas, el vector $\mathbf{F}(p)$ se encuentra en el plano tangente a la esfera con centro en el origen que pasa por P. Dé otro campo $G: \mathbb{R}^3 \to \mathbb{R}^3$ distinto de $\mathbf{F}$ que tenga esta misma propiedad.

Para cada uno de los campos $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ dados en los ejercicios 16–20, obtenga su expresión en el sistema de coordenadas cilíndricas y en el sistema de coordenadas esféricas.

16. $\mathbf{F}(x, y, z) = (x, y, z)$

17. $\mathbf{F}(x, y, z) = (x^2, xy, z)$

18. $\mathbf{F}(x, y, z) = (y, y^2, xyz)$

19. $\mathbf{F}(x, y, z) = (x + y, y + z, z)$

20. $\mathbf{F}(x, y, z) = (xy, xz, yz)$

(**) **21.** Sea $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ un campo vectorial continuo definido en el conjunto abierto U de $\mathbb{R}^2$. Una *línea de flujo* (o de campo) para $\mathbf{F}$ es un camino $\boldsymbol{\mu}: I \subseteq \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{\mu}(I) \subset U$, de clase $\mathscr{C}^1$ tal que $\boldsymbol{\mu}'(t) = \mathbf{F}(\boldsymbol{\mu}(t))$. Es decir, el camino $\boldsymbol{\mu}$ tiene por vector tangente en el punto $\mathbf{p} = \boldsymbol{\mu}(t)$ de su traza, al vector $\mathbf{F}(\mathbf{p})$. Si escribimos $\mathbf{F} = (M, N)$, en donde $M, N: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son las funciones coordenadas de F, demuestre que el camino $\boldsymbol{\mu}(t) = (x(t), y(t))$ será una línea de flujo del campo F si y solamente si

$$x'(t) = M(x(t), y(t)) \quad \text{y} \quad y'(t) = N(x(t), y(t))$$

En particular, en este ejercicio estudiaremos las líneas de flujo para campos lineales del tipo $\mathbf{F}(x, y) = (ax + by, cx + dy)$, con $ad - bc \neq 0$, para los cuales entonces el camino $\boldsymbol{\mu}(t) = (x(t), y(t))$ será una linea de flujo si y sólo si

$$x'(t) = ax(t) + by(t) \quad \text{y} \quad y'(t) = cx(t) + dy(t) \tag{S}$$

En la teoría de las ecuaciones diferenciales, se dice que el sistema (S) es un *sistema lineal homogéneo de primer orden* (de dos ecuaciones en las indeterminadas $x(t)$ y $y(t)$). Considere la matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Sea λ un valor propio de A y $\mathbf{v} \in \mathbb{R}^2$ un vector propio asociado a λ. Demuestre que

$$\boldsymbol{\mu}(t) = (x(t), y(t)) = ke^{\lambda t}\mathbf{v}$$

en donde k es una constante arbitraria, es solución del sistema (S). Más concretamente, se puede demostrar que la *solución general* del sistema (S) es de alguno de los tres tipos siguientes, atendiendo a la naturaleza de los valores propios de la matriz A:

1 Si la matriz A tiene dos valores propios reales y distintos λ_1 y λ_2, la solución general del sistema (S) es

$$\boldsymbol{\mu}(t) = (x(t), y(t)) = k_1 \exp(\lambda_1 t)\mathbf{v}_1 + k_2 \exp(\lambda_2 t)\mathbf{v}_2$$

en donde k_1 y k_2 son constantes arbitrarias, y $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios asociados a los valores propios λ_1 y λ_2 respectivamente.

2 Si la matriz A tiene dos valores propios reales e iguales, digamos $\lambda_1 = \lambda_2 = \lambda$, se tienen dos posibilidades:

2.1 Si el valor propio λ tiene asociados dos vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$ linealmente independientes, la solución general del sistema (S) es

$$\boldsymbol{\mu}(t) = (x(t), y(t)) = k_1 \exp(\lambda t)\mathbf{v}_1 + k_2 \exp(\lambda t)\mathbf{v}_2$$

en donde k_1 y k_2 son constantes arbitrarias.

2.2 Si el valor propio λ tiene solamente un vector propio $\mathbf{v}$ linealmente independiente, la solución general del sistema (S) es

$$\boldsymbol{\mu}(t) = (x(t), y(t)) = k_1 \exp(\lambda t)\mathbf{v} + k_2(t \exp(\lambda t)\mathbf{v} + \exp(\lambda t)\mathbf{u})$$

en donde k_1 y k_2 son constantes arbitrarias, y $u \in \mathbb{R}^2$ es un vector (una matriz 2×1) que satisface el sistema no homogéneo de dos ecuaciones con dos indeterminadas $(A - \lambda I)\mathbf{u} = \mathbf{v}$.

3 Si la matriz A tiene valores propios complejos, digamos $\lambda_{1,2} = \alpha \pm i\beta$, la solución general del sistema (S) es del tipo

$$\boldsymbol{\mu}(t) = (x(t), y(t)) = k_1\mathbf{w}_1(t) + k_2\mathbf{w}_2(t)$$

en donde k_1 y k_2 son constantes arbitrarias, y los vectores $\mathbf{w}_1(t)$ y $\mathbf{w}_2(t)$ son

$$\mathbf{w}_1(t) = (\mathbf{u}_1 \cos \beta t + \mathbf{u}_2 \operatorname{sen} \beta t) \exp(\alpha t)$$
$$\mathbf{w}_2(t) = (\mathbf{u}_2 \cos \beta t - \mathbf{u}_1 \operatorname{sen} \beta t) \exp(\alpha t)$$

en donde $\mathbf{u}_1 = \frac{1}{2}(\mathbf{v}+\bar{\mathbf{v}})$, $\mathbf{u}_2 = \frac{i}{2}(\mathbf{v}-\bar{\mathbf{v}})$, $\mathbf{v}$ un vector propio asociado al valor propio $\alpha+\mathbf{i}\beta$ (observe que las coordenadas de $\mathbf{v}$ son números complejos) y $\bar{\mathbf{v}}$ el vector que tiene por coordenadas los correspondientes conjugados de las coordenadas de $\mathbf{v}$. Para cada uno de los campos lineales $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ dados en los incisos siguientes, use los resultados mencionados anteriormente para obtener que las líneas de flujo de estos campos son los caminos $\boldsymbol{\mu}: \mathbb{R} \to \mathbb{R}^2$ indicados. Bosqueje algunas de las curvas (trazas de los caminos) para algunos valores de las constantes k_1 y k_2, indicando el "sentido de recorrido" de la curva (variando t de $-\infty$ a $+\infty$). El comportamiento de estas curvas cerca del origen es típico para cada uno de los casos indicados. Se proporciona también el nombre que se da al origen de coordenadas en cada caso (del que se dice, en general, que es un *punto crítico*).

a. Campo: $\mathbf{F}(x, y) = (2x + y, x + 2y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = (k_1 e^t + k_2 e^{3t}, -k_1 e^t + k_2 e^{3t})$
Comportamiento típico del caso en que los valores propios de la matriz A son ambos reales positivos (distintos).
El origen es en este caso un *nodo inestable* o *fuente*.

b. Campo: $\mathbf{F}(x, y) = (-2x + y, x - 2y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = (k_1 e^{-t} + k_2 e^{-3t}, k_1 e^{-t} - k_2 e^{-3t})$
Comportamiento típico del caso en que los valores propios de la matriz A son ambos reales negativos (distintos).
El origen es en este caso un *nodo estable* o *atractor.*

c. Campo: $\mathbf{F}(x, y) = (-2x - 2y, 3x + 5y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = (2k_1 e^{-t} + k_2 e^{4t}, -k_1 e^{-t} - 3k_2 e^{4t})$
Comportamiento típico del caso en que los valores propios de la matriz A son reales de signos opuestos.
El origen es en este caso una *silla*.

d. Campo: $\mathbf{F}(x, y) = (-2x, -2y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = (k_1 e^{-2t}, k_2 e^{-2t})$
Comportamiento típico del caso en que la matriz A tiene un solo valor propio negativo, el cual tiene dos vectores propios asociados linealmente independientes.
El origen es en este caso un *nodo propio asintóticamente estable.*

e. Campo: $\mathbf{F}(x, y) = (2x, 2y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = (k_1 e^{2t}, k_2 e^{2t})$
Comportamiento típico del caso en que la matriz A tiene un solo valor propio positivo, el cual tiene dos vectores propios asociados linealmente independientes.
El origen es en este caso un *nodo propio asintóticamente inestable.*

f. Campo: $\mathbf{F}(x, y) = (-x, 2x - y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = \big(0.5k_2 e^{-t}, (k_1 + k_2(t + 1))e^{-t}\big)$
Comportamiento típico del caso en que la matriz A tiene un solo valor propio negativo, el cual tiene solamente un vector propio linealmente independiente.
El origen es en este caso un *nodo impropio asintóticamente estable.*

g. Campo: $\mathbf{F}(x, y) = (x, 2x + y)$
Líneas de flujo: $\boldsymbol{\mu}(t) = \big(0.5k_2 e^{t}, (k_1 + k_2(t + 1))e^{t}\big)$

Comportamiento típico del caso en que la matriz A tiene un solo valor propio positivo, el cual tiene solamente un vector propio linealmente independiente.
El origen es en este caso un *nodo impropio asintóticamente inestable.*

h. Campo: $\mathbf{F}(x, y) = (2x + 5y, -x - 2y)$
Líneas de flujo: $\boldsymbol{\mu}(t) == \big(k_1 \cos t - k_2 \operatorname{sen} t, \frac{1}{5}((2k_2 - k_1) \operatorname{sen} t - (2k_1 + k_2) \cos t)\big)$
Comportamiento típico del caso en que la matriz A tiene sus valores propios complejos, con parte real igual a cero.
El origen es en este caso un *centro.*

i. Campo: $F(x, y) = (2x - y, 2x)$
Líneas de flujo: $\boldsymbol{\mu}(t) = \big(e^t(k_1 \cos t - k_2 \operatorname{sen} t), e^t((k_1 + k_2) \cos t + (k_1 - k_2) \operatorname{sen} t)\big)$
Comportamiento típico del caso en que la matriz A tiene sus valores propios complejos, con la parte real positiva.
El origen es en este caso un *foco inestable.*

j. Campo: $F(x, y) = (-2x - y, 2x)$
Líneas de flujo: $\boldsymbol{\mu}(t) = \big(e^{-t}(k_1 \cos t - k_2 \operatorname{sen} t), e^{-t}((k_1 + k_2) \operatorname{sen} t - (k_1 - k_2) \cos t)\big)$
Comportamiento típico del caso en que la matriz A tiene sus valores propios complejos, con la parte real negativa.
El origen es en este caso un *foco estable.*

7.3 Integrales de línea

Comencemos por dar la definición de las integrales que estudiaremos en esta sección.

Definición. Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ un campo vectorial continuo y sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ un camino de clase $\mathscr{C}^1$ cuya traza está contenida en U, es decir, $\boldsymbol{\lambda}([a, b]) \subset U$. La integral de línea del campo F a lo largo del (o sobre el) camino $\boldsymbol{\lambda}$, se define denotada por

$$\int_{\boldsymbol{\lambda}} \mathbf{F}, \quad \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}, \quad \text{o} \quad \int_{\boldsymbol{\lambda}} F_1(x)\, dx_1 + F_2(x)\, dx_2 + \cdots + F_n(x)\, dx_n$$

como

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\, dt$$

(donde $\cdot$ es el producto punto) o bien, de modo más explícito como

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_a^b (F_1(\boldsymbol{\lambda}(t))\lambda_1'(t) + F_2(\boldsymbol{\lambda}(t))\lambda_2'(t) + \cdots + F_n(\boldsymbol{\lambda}(t))\lambda_n'(t))\, dt \\ &= \int_a^b \sum_{i=1}^n F_i(\boldsymbol{\lambda}(t))\lambda_i'(t)\, dt = \sum_{i=1}^n \int_a^b F_i(\boldsymbol{\lambda}(t))\lambda_i'(t)\, dt\end{aligned}$$ ■

Nótese entonces que para calcular la integral del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}$, evaluamos la función $\mathbf{F}$ en los puntos (imágenes) del camino (es decir, en los puntos de la curva $\boldsymbol{\lambda}$. Esto es posible, pues hemos pedido que las imágenes de $\boldsymbol{\lambda}$ se encuentren dentro del dominio del campo $\mathbf{F}$) y

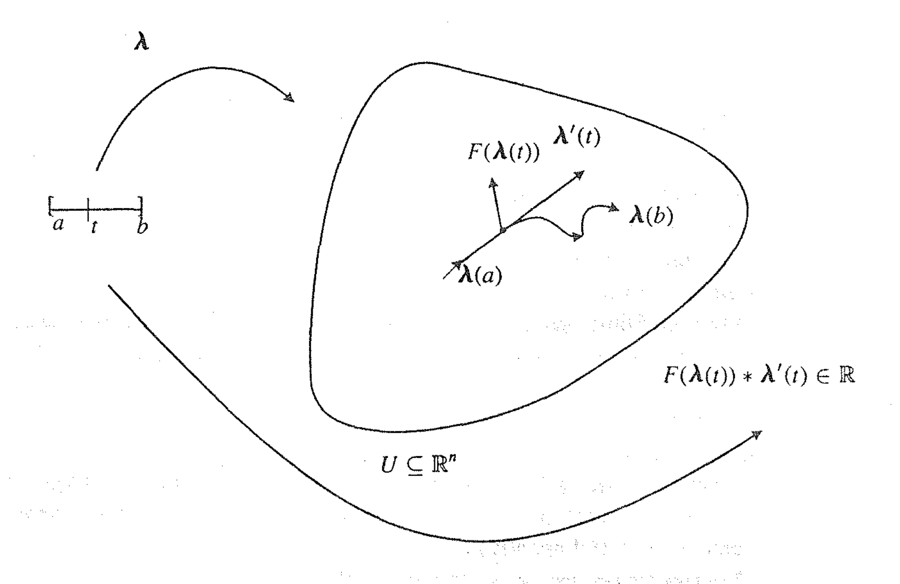


Figura 1. La integral del campo **F** a lo largo del camino **λ**.

hacemos el producto punto con el vector velocidad $\boldsymbol{\lambda}'(t)$ del camino $\boldsymbol{\lambda}$. El resultado es una función real de la variable real t, que integramos entre a y b (donde se define $\boldsymbol{\lambda}$).

Cuando el camino $\boldsymbol{\lambda}$ es cerrado, se suele usar la notación

$$\oint_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda} \quad \text{o} \quad \oint_{\boldsymbol{\lambda}} F_1(x)\,dx_1 + F_2(x)\,dx_2 + \cdots + F_n(x)\,dx_n$$

para indicar la integral de línea del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$. Nótese que siendo continuo el campo F, la existencia de la integral de línea de F a lo largo de cualquier camino de clase $\mathscr{C}^1$ está asegurada, pues en tal caso el integrando $\mathbf{F}(\boldsymbol{\lambda}(t))\cdot\boldsymbol{\lambda}'(t)$ es una función continua en el intervalo $[a, b]$ y por lo tanto su integral en tal intervalo existe.

En este capítulo tomaremos como ejemplo sobre todo campos en $\mathbb{R}^2$ o $\mathbb{R}^3$. Por ejemplo, el campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ lo escribimos como $\mathbf{F}(x, y) = (F_1(x, y), F_2(x, y))$, y el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ (tal que $\boldsymbol{\lambda}([a, b]) \subset U$) como $\boldsymbol{\lambda}(t) = (x(t), y(t))$, de modo que la integral de F a lo largo de $\boldsymbol{\lambda}$ se ve en este caso como

$$\int_{\boldsymbol{\lambda}} F_1(x, y)\,dx + F_2(x, y)\,dy = \int_a^b \big(F_1(x(t), y(t))x'(t) + F_2(x(t), y(t))y'(t)\big)\,dt$$

Del mismo modo, escribimos el campo en $\mathbb{R}^3$, $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ como $\mathbf{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z))$, y el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ (de modo que $\boldsymbol{\lambda}([a, b]) \subset U$) como $\boldsymbol{\lambda}(t) = (x(t), y(t), z(t))$, para que en este caso la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ se ve como

$$\int_{\boldsymbol{\lambda}} F_1(x, y, z)\,dx + F_2(x, y, z)\,dy + F_3(x, y, z)\,dz$$
$$= \int_a^b \big(F_1(x(t), y(t), z(t))x'(t) + F_2(x(t), y(t), z(t))y'(t) + F_3(x(t), y(t), z(t))z'(t)\big)\,dt$$

Veamos algunos ejemplos.

Ejemplo 1. Sea $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ el campo $\mathbf{F}(x, y) = (x + y, y)$ y sea $\boldsymbol{\lambda}\colon [0, 1] \to \mathbb{R}^2$ el camino $\boldsymbol{\lambda}(t) = (t, t^2)$. La integral de F a lo largo de $\boldsymbol{\lambda}$ es

$$\int_{\boldsymbol{\lambda}} F_1(x, y)\,dx + F_2(x, y)\,dy = \int_0^1 \left((t + t^2)(t)' + (t^2)(t^2)'\right) dt$$
$$= \int_0^1 (t + t^2 + 2t^3)\,dt = \frac{1}{2} + \frac{1}{3} + \frac{1}{2} = \frac{4}{3}$$ ■

Ejemplo 2. Consideremos el mismo campo del ejemplo anterior, $\mathbf{F}(x, y) = (x + y, y)$ y tomemos el camino $\boldsymbol{\mu}\colon [0, k^{-1}] \to \mathbb{R}^2$ dado por $\boldsymbol{\mu}(t) = (kt, k^2t^2)$, donde k es un número positivo. Obsérvese que este nuevo camino es una reparametrización del camino $\boldsymbol{\lambda}$ considerado en el ejemplo anterior (de hecho se tiene $\boldsymbol{\mu} = \boldsymbol{\lambda} \circ \phi$, donde $\phi\colon [0, 1] \to [0, 1]$ es $\phi(t) = kt$). Calculemos la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\mu}$. Esta es

$$\int_{\boldsymbol{\mu}} F_1(x, y)\,dx + F_2(x, y)\,dy = \int_0^{k^{-1}} \left((kt + k^2t^2)(kt)' + k^2t^2(k^2t^2)'\right) dt$$
$$= \int_0^{k^{-1}} (k^2t + k^3t^2 + 2k^4t^3)\,dt$$
$$= \frac{1}{2}k^2(k^{-1})^2 + \frac{1}{3}k^3(k^{-1})^3 + \frac{1}{2}k^4(k^{-1})^4 = \frac{4}{3}$$

que es el mismo resultado obtenido en el ejemplo anterior. Esta es una propiedad general de las integrales de línea que estudiaremos en esta sección. Establece que una integral de línea es invariante por reparametrizaciones del camino sobre el que se integra el campo $\mathbf{F}$. ■

Ejemplo 3. Tomemos nuevamente el campo de los dos ejemplos anteriores $\mathbf{F}(x, y) = (x + y, y)$ e integrémoslo ahora sobre el camino inverso al camino dado en el ejemplo 1. Este es $-\boldsymbol{\lambda}\colon [0, 1] \to \mathbb{R}^2$, $-\boldsymbol{\lambda}(t) = (1 - t, (1 - t)^2)$. La integral de F a lo largo de $-\boldsymbol{\lambda}$ es

$$\int_{-\boldsymbol{\lambda}} F_1(x, y)dx + F_2(x, y)\,dy = \int_0^1 \left([(1 - t) + (1 - t)^2](1 - t)' + (1 - t)^2((1 - t)^2)'\right) dt$$
$$= \int_0^1 \left(-(1 - t) - (1 - t)^2 - 2(1 - t)^3\right) dt$$
$$= \frac{1}{2}(1 - t)^2 + \frac{1}{3}(1 - t)^3 + \frac{1}{2}(1 - t)^4\Big]_0^1$$
$$= -\frac{1}{2} - \frac{1}{3} - \frac{1}{2} = -\frac{4}{3}$$

Nótese que

$$\int_{-\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = -\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

Esta será otra de las propiedades de las integrales de línea que estudiaremos: la integral de línea de un campo $\mathbf{F}$ sobre el camino inverso $-\boldsymbol{\lambda}$ es el negativo de la integral de línea del campo F sobre $\boldsymbol{\lambda}$. ■

Ejemplo 4. Con el mismo campo $\mathbf{F}$ de los ejemplos anteriores $\mathbf{F}(x, y) = (x + y, y)$, hagamos ahora la integral de línea de $\mathbf{F}$ a lo largo del camino $\boldsymbol{\nu}$: $[0, 1] \to \mathbb{R}^2$, $\boldsymbol{\nu}(t) = (t, t^3)$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\nu}} F_1(x, y)\,dx + F_2(x, y)\,dy &= \int_0^1 \left((t + t^3)(t)' + (t^3)(t^3)'\right) dt \\ &= \int_0^1 (t + t^3 + 3t^5)\,dt = \frac{1}{2} + \frac{1}{4} + \frac{1}{2} = \frac{5}{4}\end{aligned}$$ ■

Préstese atención a los resultados obtenidos en el ejemplo anterior y en el ejemplo 1. Obsérvese que los caminos $\boldsymbol{\lambda}$ (del ejemplo 1) y $\boldsymbol{\nu}$ (del ejemplo 4) comparten su punto inicial $\boldsymbol{\lambda}(0) = \boldsymbol{\nu}(0) = (0, 0)$ y su punto final $\boldsymbol{\lambda}(1) = \boldsymbol{\nu}(1) = (1, 1)$. El camino $\boldsymbol{\lambda}$ recorre el arco de parábola $y = x^2$, $0 \leq x \leq 1$, y el camino $\boldsymbol{\nu}$ recorre el arco de parábola cúbica $y = x^3$, $0 \leq x \leq 1$. Sin embargo, los resultados de las integrales son diferentes. Esto hace suponer que el valor de la integral de línea de un campo $\mathbf{F}$ sobre un camino dado, no depende solamente de los puntos inicial y final del camino, sino de la función misma que lo define. Este es un asunto muy importante que se expondrá ampliamente en la siguiente sección. Por el momento, queremos sembrar la inquietud al respecto con el siguiente par de ejemplos.

Ejemplo 5. Consideremos ahora el campo $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y^2, 2xy)$ y hagamos la integral de línea de este campo sobre el camino del ejemplo 1, $\boldsymbol{\lambda}$: $[0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\lambda}} F_1(x, y)dx + F_2(x, y)dy &= \int_0^1 \left((t + t^4)(t)' + 2(t)(t^2)(t^2)'\right)dt \\ \int_0^1 (t + 5t^4)dt &= \frac{1}{2} + 1 = \frac{3}{2}\end{aligned}$$ ■

Ejemplo 6. Con el mismo campo del ejemplo anterior $\mathbf{F}(x, y) = (x + y^2, 2xy)$, hagamos la integral de línea sobre el camino del ejemplo 4, $\boldsymbol{\nu}$: $[0, 1] \to \mathbb{R}^2$, $\boldsymbol{\nu}(t) = (t, t^3)$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\nu}} F_1(x, y)\,dx + F_2(x, y)\,dy &= \int_0^1 \left((t + t^6)(t)' + 2(t)(t^3)(t^3)'\right) dt \\ \int_0^1 (t + 7t^6)\,dt &= \frac{1}{2} + 1 = \frac{3}{2}\end{aligned}$$ ■

Nuevamente llamamos la atención sobre el hecho de que en este par de ejemplos se obtuvo el mismo resultado al integrar el campo $\mathbf{F}$ con dos caminos distintos que comparten su punto inicial y su punto final. La pregunta previa al ejemplo 5 sigue en el aire: ¿la integral de línea de un campo $\mathbf{F}$ sobre un camino $\boldsymbol{\lambda}$: $[a, b] \to \mathbb{R}^2$ depende solamente de los puntos inicial $\boldsymbol{\lambda}(a)$ y final $\boldsymbol{\lambda}(b)$ del camino $\boldsymbol{\lambda}$ (como en el caso de los ejemplos 5 y 6), o depende también de la función misma que define al camino (como en el caso de los ejemplos 1 y 4)? Este es uno de los temas centrales a tratar en la próxima sección de este capítulo.

Después de estos ejemplos preliminares, en los que se han calculado algunas integrales de línea y anunciado ya algunas de las propiedades de éstas, estudiaremos de modo formal dichas propiedades, las cuales nos descubrirán el comportamiento de las integrales objeto de estudio en este capítulo. Ya que las integrales de línea se calculan finalmente en términos de integrales "simples" de funciones

reales de variable real, es de esperarse que la propiedad de linealidad de estas últimas se herede a las integrales de línea. En efecto, si $\mathbf{F}, \mathbf{G}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ son dos campos continuos y $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, $\boldsymbol{\lambda}([a, b]) \subset U$, un camino de clase $\mathscr{C}^1$, se tiene que

$$\int_{\boldsymbol{\lambda}} (\mathbf{F} + c\mathbf{G}) \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} + c \int_{\boldsymbol{\lambda}} \mathbf{G} \cdot d\boldsymbol{\lambda}$$

donde $c \in \mathbb{R}$, y el campo $\mathbf{F} + c\mathbf{G}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, es $(\mathbf{F} + c\mathbf{G})(\mathbf{x}) = \mathbf{F}(\mathbf{x}) + c\mathbf{G}(\mathbf{x})$, $\mathbf{x} \in U$. Dejamos al lector que verifique esta propiedad. En el siguiente teorema se resumen las otras propiedades (menos obvias que la linealidad) que poseen las integrales de línea.

Teorema 7.3.1 Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un campo vectorial y $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$ tal que $\boldsymbol{\lambda}([a, b]) \subset U$. Supongamos que $\mathbf{F}$ es continuo (con lo cual la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ existe).

a. Sea $\phi: [c, d] \to [a, b]$ una función sobreyectiva de clase $\mathscr{C}^1$, y sea $\boldsymbol{\mu}: [c, d] \to \mathbb{R}^n$ el camino $\boldsymbol{\mu} = \boldsymbol{\lambda} \circ \phi$.

a1. Si $\phi(c) = a$ y $\phi(d) = b$, entonces

$$\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

a2. Si $\phi(c) = b$ y $\phi(d) = a$, entonces

$$\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} = -\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

En particular, si $\boldsymbol{\mu}$ es una reparametrización de $\boldsymbol{\lambda}$, entonces

a1. $\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ si $\boldsymbol{\mu}$ conserva la orientación de $\boldsymbol{\lambda}$.

a2. $\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} = -\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ si $\boldsymbol{\mu}$ invierte la orientación de $\boldsymbol{\lambda}$. En particular se tiene $\int_{-\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\mu} = -\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$.

b. Si $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, entonces

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} \mathbf{F} \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_2} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

Demostración.

a. Usaremos la fórmula de cambio de variable en la integral de Riemann: Si $F: [a, b] \to \mathbb{R}$ es continua, y $G: [c, d] \to \mathbb{R}$ es de clase $\mathscr{C}^1$ y de modo que $G([c, d]) \subseteq [a, b]$, entonces

$$\int_{G(c)}^{G(d)} F(x)\, dx = \int_c^d F(G(t))G'(t)\, dt$$

a1. Supongamos que $\phi(c) = a$, $\phi(d) = b$. Así se tiene

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda} &= \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t))\cdot\boldsymbol{\lambda}'(t)dt \overset{t=\phi(s)}{=\!=\!\longrightarrow} \\ &= \int_c^d \mathbf{F}(\boldsymbol{\lambda}(\phi(s)))\cdot\boldsymbol{\lambda}'(\phi(s))\phi'(s)\,ds = \int_c^d \mathbf{F}((\boldsymbol{\lambda}\circ\phi)(s))\cdot(\boldsymbol{\lambda}\circ\phi)'(s)\,ds \\ &= \int_{\boldsymbol{\lambda}\circ\phi} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\mu}} \mathbf{F}\cdot d\boldsymbol{\lambda}\end{aligned}$$

a2. De igual modo, si $\phi(c) = b$, $\phi(d) = a$, se tiene

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda} &= \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t))\cdot\boldsymbol{\lambda}'(t)\,dt \overset{t=\phi(s)}{=\!=\!\longrightarrow} \\ &= \int_d^c \mathbf{F}(\boldsymbol{\lambda}(\phi(s)))\cdot\boldsymbol{\lambda}'(\phi(s))\phi'(s)\,ds = -\int_c^d \mathbf{F}((\boldsymbol{\lambda}\circ\phi)(s))\cdot(\boldsymbol{\lambda}\circ\phi)'(s)\,ds \\ &= -\int_{\boldsymbol{\lambda}\circ\phi} \mathbf{F}\cdot d\boldsymbol{\lambda} = -\int_{\boldsymbol{\mu}} \mathbf{F}\cdot d\boldsymbol{\lambda}\end{aligned}$$

b. Supongamos que $\boldsymbol{\lambda}_1$ está definido en $[a, c]$ y $\boldsymbol{\lambda}_2$ en $[c, b]$. Entonces $\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda}(t) = \begin{cases} \boldsymbol{\lambda}_1(t) & \text{si } t \in [a, c] \\ \boldsymbol{\lambda}_2(t) & \text{si } t \in (c, b] \end{cases}$$

y la integral de $\mathbf{F}$ sobre $\boldsymbol{\lambda}$ es

$$\begin{aligned}\int \boldsymbol{\lambda}\mathbf{F}\cdot d\boldsymbol{\lambda} &= \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t))\cdot\boldsymbol{\lambda}'(t)\,dt \\ &= \int_a^c \mathbf{F}(\boldsymbol{\lambda}(t))\cdot\boldsymbol{\lambda}'(t)\,dt + \int_c^b \mathbf{F}(\boldsymbol{\lambda}(t))\cdot\boldsymbol{\lambda}'(t)\,dt = \int_{\boldsymbol{\lambda}_1} F\cdot d\boldsymbol{\lambda} + \int \boldsymbol{\lambda}_2\mathbf{F}\cdot d\boldsymbol{\lambda} \quad \text{Q.E.D.}\end{aligned}$$

La propiedad **b** del teorema anterior nos permite definir la integral de un campo $\mathbf{F}$ sobre un camino $\boldsymbol{\lambda}$ seccionalmente $\mathscr{C}^1$. En este caso podemos escribir $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \cdots + \boldsymbol{\lambda}_k$ donde $\boldsymbol{\lambda}_i$ es un camino de clase $\mathscr{C}^1$, $i = 1, 2, \ldots, k$. Entonces definimos la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ como

$$\int_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda} = \sum_{i=1}^{k} \int_{\boldsymbol{\lambda}_i} \mathbf{F}\cdot d\boldsymbol{\lambda}$$

Por otra parte, una consecuencia muy importante de la propiedad del inciso a del teorema anterior (propiedad que se ilustra con los ejemplos 1 y 2 vistos anteriormente) es que para calcular una integral de línea de un campo $\mathbf{F}$ sobre un camino $\boldsymbol{\lambda}$ es suficiente con tener la información de *la traza del camino* (con la información correspondiente de dónde comienza, dónde termina, y cómo se recorre tal traza) para poder calcular la integral. En efecto, basta conseguir UN CAMINO (CUALQUIERA) concreto $\boldsymbol{\lambda}$ que tenga la traza dada para poder hacer el cálculo de la integral: cualquier reparametrización de $\boldsymbol{\lambda}$ producirá el mismo valor de la integral buscada. Entonces tiene sentido hablar de la integral de un campo $\mathbf{F}$ sobre *una curva* $\boldsymbol{\lambda}$ (el conjunto de imágenes del camino $\boldsymbol{\lambda}$).

Más aún, la propiedad del inciso a acepta composiciónes del camino $\boldsymbol{\lambda}$ (a lo largo del cual se calcula la integral) con funciones ϕ tales que $\boldsymbol{\mu} = \boldsymbol{\lambda} \circ \phi$ *no es una reparametrización de* $\boldsymbol{\lambda}$ (lo único que se pide es que ϕ sea sobreyectiva y que los puntos extremos de $\boldsymbol{\mu}$ coincidan con los correspondientes de $\boldsymbol{\lambda}$). Esto significa que el nuevo camino $\boldsymbol{\mu}$ que recorre la traza de $\boldsymbol{\lambda}$, puede hacerlo no sólo con velocidad distinta y con diferente orientación, sino que éste puede "sufrir regresos múltiples" en su trayecto (fenómeno que está prohibido en una reparametrización de un camino, según se expuso en el capítulo 5), y aún así la integral se mantiene invariable (cambiando de signo en el caso de que se inviertan los puntos inicial y final del camino). Este hecho descubre que el valor de una integral de línea de un campo $\mathbf{F}$ *no es acumulativo,* en contraste, por ejemplo, con el valor de la longitud de la curva descrita por el camino (y en contraste también con las integrales de línea respecto de la longitud de arco que estudiaremos en la sección 7 de este capítulo). Por ejemplo, consideremos la integral de línea calculada en el ejemplo 1, con el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y, y)$ y el camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$. En el ejemplo 2 se calculó nuevamente la integral con reparametrizaciones de $\boldsymbol{\lambda}$ que conservan la orientación. Sea $\phi: [-\pi/2, 5\pi/2] \to [0, 1]$ la función $\phi(s) = 0.5(\operatorname{sen} s + 1)$. Esta es una función sobreyectiva de modo que $\phi(-\pi/2) = 0$, $\phi(5\pi/2) = 1$, la cual "recorre su imagen tres veces" (ver figura 2). La composición $\boldsymbol{\mu} = \boldsymbol{\lambda} \circ \phi: [-\pi/2, 5\pi/2] \to \mathbb{R}^2$, $\boldsymbol{\mu}(s) = \big(0.5(\operatorname{sen} s + 1), (0.5(\operatorname{sen} s + 1))^2\big)$ es un camino (que no es una reparametrización de $\boldsymbol{\lambda}$) que parte del punto (0, 0), va hacia el punto (1, 1) por la parábola $y = x^2$, regresa al (0, 0) y, finalmente, va de nuevo al (1, 1).

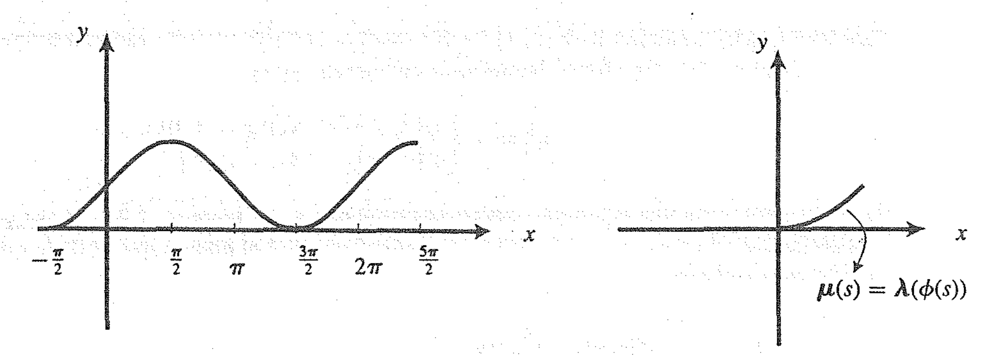


Figura 2. La función $\phi(s) = 0.5(\operatorname{sen} s + 1)$ y el camino $\boldsymbol{\mu}(s) = \boldsymbol{\lambda}(\phi(s))$

La integral del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\mu}$ es

$$\begin{aligned}
\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} &= \int_{-\pi/2}^{5\pi/2} \big[\big(0.5(\operatorname{sen} s + 1) + (0.5(\operatorname{sen} s + 1))^2\big)(0.5 \cos s) \\
&\quad + \big((0.5(\operatorname{sen} s + 1))^2\big)(\operatorname{sen} s + 1)(0.5 \cos s)\big]\, ds \\
&= \int_{-\pi/2}^{5\pi/2} \Big(\frac{1}{8} \operatorname{sen}^3 s + \frac{1}{2} \operatorname{sen}^2 s + \frac{7}{8} \operatorname{sen} s + \frac{1}{2}\Big) \cos s\, ds \\
&= \left[\frac{1}{32} \operatorname{sen}^4 s + \frac{1}{6} \operatorname{sen}^3 s + \frac{7}{16} \operatorname{sen}^2 s + \frac{1}{2} \operatorname{sen} s\right]_{-\pi/2}^{5\pi/2} \\
&= \left(\frac{1}{32} + \frac{1}{6} + \frac{7}{16} + \frac{1}{2}\right) - \left(\frac{1}{32} - \frac{1}{6} + \frac{7}{16} - \frac{1}{2}\right) = \frac{4}{3}
\end{aligned}$$

que es el mismo valor obtenido en los ejemplos 1 y 2.

Veamos más ejemplos.

Ejemplo 7. Sea $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ el campo $\mathbf{F}(x, y) = (xy, x^2y)$. Consideremos el camino $\boldsymbol{\lambda}: [-1, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, |t|)$. Este no es un camino de clase $\mathscr{C}^1$ (no es ni siquiera diferenciable), pero es seccionalmente $\mathscr{C}^1$, pues $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, donde $\boldsymbol{\lambda}_1: [-1, 0] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_1(t) = (t, -t)$, $\boldsymbol{\lambda}_2: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_2(t) = (t, t)$, que son caminos de clase $\mathscr{C}^1$. Entonces la integral del campo $\mathbf{F}$ sobre $\boldsymbol{\lambda}$ es

$$\begin{aligned}
\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_{\boldsymbol{\lambda}_1} \mathbf{F} \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_2} F \cdot d\boldsymbol{\lambda} \\
&= \int_{-1}^{0} \left((t)(-t)(t)' + (t^2)(-t)(-t)'\right) dt + \int_{0}^{1} \left((t)(t)(t)' + (t^2)(t)(t)'\right) dt \\
&= \int_{-1}^{0} (-t^2 + t^3) dt + \int_{0}^{1} (t^2 + t^3) dt \\
&= \left[-\frac{1}{3}t^3 + \frac{1}{4}t^4\right]_{-1}^{0} + \left[\frac{1}{3}t^3 + \frac{1}{4}t^4\right]_{0}^{1} \\
&= -\frac{1}{3} - \frac{1}{4} + \frac{1}{3} + \frac{1}{4} = 0
\end{aligned}$$

Obsérvese que el camino $\boldsymbol{\mu}: [-1, 1] \to \mathbb{R}^2$, $\boldsymbol{\mu}(t) = (t^3, t^2|t|)$ es una reparametrización del camino $\boldsymbol{\lambda}$, y es de clase $\mathscr{C}^1$. En efecto, la derivada del camino $\boldsymbol{\mu}$ es:

$$\boldsymbol{\mu}'(t) = \begin{cases} (3t^2, -3t^2) & \text{si } t \in [-1, 0] \\ (3t^2, 3t^2) & \text{si } t \in (0, 1] \end{cases}$$

la cual es una función continua. Según la propiedad **a** del teorema 7.2.1, la integral del campo $\mathbf{F}$ puede ser calculada sobre el camino $\boldsymbol{\mu}$, y el resultado será el mismo que el de la integral de $\mathbf{F}$ sobre $\boldsymbol{\lambda}$. Se tiene entonces

$$\begin{aligned}
\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} &= \int_{-1}^{1} \mathbf{F}(\boldsymbol{\mu}(t)) \cdot \boldsymbol{\mu}'(t)\, dt \\
&= \int_{-1}^{0} \mathbf{F}(\boldsymbol{\mu}(t)) \cdot \boldsymbol{\mu}'(t)\, dt + \int_{0}^{1} \mathbf{F}(\boldsymbol{\mu}(t)) \cdot \boldsymbol{\mu}'(t)\, dt \\
&= \int_{-1}^{0} \left((t^3)(-t^3)(3t^2) + (t^6)(-t^3)(-3t^2)\right) dt + \int_{0}^{1} \left((t^3)(t^3)(3t^2) + (t^6)(t^3)(3t^2)\right) dt \\
&= \int_{-1}^{0} (-3t^8 + 3t^{11})\, dt + \int_{0}^{1} (3t^8 + 3t^{11})\, dt \\
&= \left[-\frac{3}{9}t^9 + \frac{3}{12}t^{12}\right]_{-1}^{0} + \left[\frac{3}{9}t^9 + \frac{3}{12}t^{12}\right]_{0}^{1} \\
&= -\frac{1}{3} - \frac{1}{4} + \frac{1}{3} + \frac{1}{4} = 0
\end{aligned}$$

■

Ejemplo 8. Se quiere calcular la integral de línea del campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ $\mathbf{F}(x, y) = (2x + y, -x + 4xy)$ sobre el círculo $x^2 + y^2 = 1$ recorrido (una vez) en sentido antihorario (contrario a las manecillas

del reloj). Obsérvese que podemos escribir la integral por calcular de la siguiente manera

$$\oint_{x^2+y^2=1} F_1(x, y)\, dx + F_2(x, y)\, dy$$

donde se indica que la integral es sobre un camino cerrado, y se da la información de *la curva* sobre la que se va a realizar la integral. Debemos entonces tomar una parametrización (un camino) de la curva $C = \{(x, y)|x^2 + y^2 = 1\}$. Ésta puede ser, por ejemplo, $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t)$ —de hecho, puede ser cualquier otra que al lector se le ocurra y/o se le antoje—. Se tiene entonces

$$\begin{aligned}\oint_{x^2+y^2=1} F_1(x, y)dx + F_2(x, y)\, dy &= \int_0^{2\pi} \big((2\cos t + \operatorname{sen} t)(-\operatorname{sen} t) + (-\cos t + 4\cos t \operatorname{sen} t)(\cos t)\big)\, dt \\ &= \int_0^{2\pi} (-2\operatorname{sen} t \cos t - 1 + 4\cos^2 t \operatorname{sen} t)\, dt \\ &= \left[-\operatorname{sen}^2 t - t - \frac{4}{3}\cos^3 t\right]_0^{2\pi} = -2\pi\end{aligned}$$

■

Ejemplo 9. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + 4y, ax + y)$ donde a es una constante. Hagamos la integral de línea de este campo sobre la curva $C = \{(x, y)\big||x| + |y| = 1\}$. Es fácil ver que esta curva es un cuadrado con vértices en los puntos $(\pm 1, 0)$ y $(0, \pm 1)$. En efecto:

a Para $x \geq 0$, $y \geq 0$, se tiene la recta $C_1: x + y = 1$.

b Para $x \leq 0$, $y \geq 0$, se tiene la recta $C_2: -x + y = 1$.

c Para $x \leq 0$, $y \leq 0$, se tiene la recta $C_3: -x - y = 1$.

d Para $x \geq 0$, $y \leq 0$, se tiene la recta $C_4: x - y = 1$.

Parametrizando cada una de estos segmentos de recta, tenemos

a un camino para $C_1: x + y = 1$, es $\boldsymbol{\lambda}_1: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_1(t) = (1 - t, t)$.

b un camino para $C_2: -x + y = 1$, es $\boldsymbol{\lambda}_2: [-1, 0] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_2(t) = (-1 - t, -t)$.

c un camino para $C_3: -x - y = 1$, es $\boldsymbol{\lambda}_3: [-1, 0] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_3(t) = (t, -1 - t)$.

d un camino para $C_4: x - y = 1$, es $\boldsymbol{\lambda}_4: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_4(t) = (t, t - 1)$.

Nótese entonces que el camino (seccionalmente $\mathscr{C}^1$) $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3 + \boldsymbol{\lambda}_4$ es una parametrización de la curva C. Se tiene entonces

$$\oint_C F_1(x, y)\, dx + F_2(x, y)\, dy = \oint_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \sum_{i=1}^{4} \int_{\boldsymbol{\lambda}_i} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

Hagamos cada una de las integrales indicadas

$$\begin{aligned}\int_{\boldsymbol{\lambda}_1} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_0^1 \big((1 - t + 4t)(1 - t)' + (a(1 - t) + t)(t)'\big)\, dt \\ &= \int_0^1 (a - 1 - (a + 2)t)\, dt = \left[(a - 1)t - \frac{1}{2}(a + 2)t^2\right]_0^1 = \frac{1}{2}a - 2\end{aligned}$$

$$\int_{\lambda_2} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{-1}^{0} \left((-1-t-4t)(-1-t)' + (a(-1-t)-t)(-t)'\right) dt$$
$$= \int_{-1}^{0} ((a+6)t + a + 1)\, dt = \left[\frac{1}{2}(a+6)t^2 + (a+1)t\right]_{-1}^{0} = \frac{1}{2}a - 2$$
$$\int_{\lambda_3} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{-1}^{0} \left((t+4(-1-t))(t)' + (at-1-t)(-1-t)'\right) dt$$
$$= \int_{-1}^{0} (-(a+2)t - 3)\, dt = \left[-\frac{1}{2}(a+2)t^2 - 3t\right]_{-1}^{0} = \frac{1}{2}a - 2$$
$$\int_{\lambda_4} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{0}^{1} \left((t+4(t-1))(t)' + (at+t-1)(t-1)'\right) dt$$
$$= \int_{0}^{1} ((a+6)t - 5)\, dt = \left[\frac{1}{2}(a+6)t^2 - 5t\right]_{0}^{1} = \frac{1}{2}a - 2$$

Entonces

$$\oint_{\lambda} \mathbf{F}\cdot d\boldsymbol{\lambda} = 4\left(\frac{1}{2}a - 2\right) = 2a - 8$$ ■

Ejemplo 10. Consideremos nuevamente el campo $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ del ejemplo anterior, $\mathbf{F}(x, y) = (x+4y, ax+y)$ y hagamos ahora la integral de $\mathbf{F}$ a lo largo del círculo $x^2+y^2 = r^2$, recorrido una vez en sentido antihorario. Como hicimos en el ejemplo 8, podemos tomar el camino $\boldsymbol{\lambda}\colon [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (r\cos t, r\operatorname{sen} t)$ como parametrización del círculo. Así

$$\oint_{x^2+y^2=r^2} F_1(x, y)\, dx + F_2(x, y)\, dy = \int_0^{2\pi} \left((r\cos t + 4r\operatorname{sen} t)(-r\operatorname{sen} t) + (ar\cos t + r\operatorname{sen} t)(r\cos t)\right) dt$$
$$= \int_0^{2\pi} (ar^2\cos^2 t - 4r^2\operatorname{sen}^2 t)\, dt = \pi r^2(a-4)$$ ■

Ejemplo 11. Consideremos el campo $\mathbf{F}\colon \mathbb{R}^2 - \{(0,0)\} \to \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$$

y sea $\boldsymbol{\lambda}_k$ el camino $\boldsymbol{\lambda}_k\colon [0, k\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_k(t) = (r\cos t, r\operatorname{sen} t)$, donde $0 < k \leq 2$. En forma geométrica, la curva que describe $\boldsymbol{\lambda}$ es un pedazo del círculo $x^2 + y^2 = r^2$ con ángulo central $k\pi$ (figura 3).

Se quiere calcular la integral del campo $\mathbf{F}$ sobre el camino $\boldsymbol{\lambda}_k$. Se tiene

$$\int_{\lambda_k} F_1(x, y)\, dx + F_2(x, y)\, dy = \int_0^{k\pi} \left(\frac{-r\operatorname{sen} t}{r^2} + (-r\operatorname{sen} t) + \frac{r\cos t}{r^2}(r\cos t)\right) dt$$
$$= \int_0^{k\pi} dt = k\pi$$

Un hecho interesante: el resultado *no depende de r*. Esto significa que la integral de línea del campo $\mathbf{F}$ resulta ser invariante por "ampliación o reducción" del círculo, con centro en el origen sobre el cual se está integrando. Este hecho se explicará en la sección 5. ■

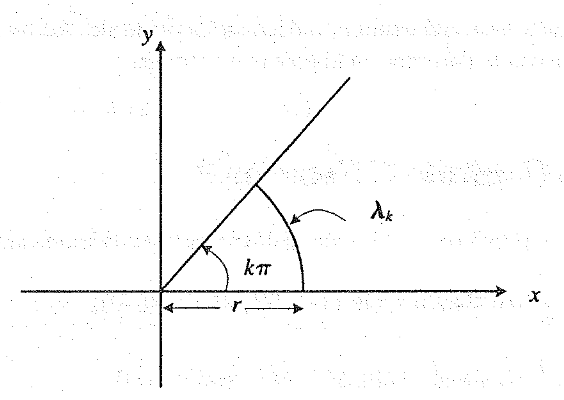


Figura 3. El camino $\boldsymbol{\lambda}_k$ del ejemplo 11.

Ejemplo 12. Hasta este momento hemos efectuado cálculos de integrales de línea con campos y caminos en $\mathbb{R}^2$. Consideremos ahora el campo en $\mathbb{R}^3$, $\mathbf{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$, dado por $\mathbf{F}(x, y, z) = (yz, xz, xy)$ y sea $\boldsymbol{\lambda}\colon [0, \pi/4] \to \mathbb{R}^3$ el camino $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t, t)$. En forma geométrica $\boldsymbol{\lambda}([0, \pi/4])$ corresponde a la octava parte de una vuelta de una hélice dibujada sobre el cilindro $x^2 + y^2 = 1$ (ver ejemplo 6 de la sección 2 del capítulo 5) comenzando en el punto $(1, 0, 0)$. La integral de línea del campo $\mathbf{F}$ a lo largo de la curva $\boldsymbol{\lambda}$ es

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_0^{\pi/4} \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\, dt \\ &= \int_0^{\pi/4} \big(t \operatorname{sen} t, t \cos t, \operatorname{sen} t \cos t\big) \cdot \big(-\operatorname{sen} t, \cos t, 1\big) dt \\ &= \int_0^{\pi/4} (-t \operatorname{sen}^2 t + t \cos^2 t + \operatorname{sen} t \cos t)\, dt \\ &= \int_0^{\pi/4} (t \cos 2t + \operatorname{sen} t \cos t)\, dt = \Big[t \operatorname{sen} t \cos t\Big]_0^{\pi/4} = \frac{\pi}{8}\end{aligned}$$

■

Ejemplo 13. Consideremos nuevamente el campo del ejemplo anterior $\mathbf{F}(x, y, z) = (yz, xz, xy)$ y tomemos ahora un camino $\boldsymbol{\mu}$ cuya traza sea una línea recta (en $\mathbb{R}^3$) que una al punto $(1, 0, 0)$ con el punto $(\sqrt{2}/2, \sqrt{2}/2\, \pi/4)$. Obsérvese entonces que en tal caso el camino $\boldsymbol{\mu}$ comparte sus puntos inicial y final con el camino $\boldsymbol{\lambda}$ del ejemplo anterior. Este camino puede ser, por ejemplo, $\boldsymbol{\mu}\colon [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\mu}(t) = (\sqrt{2}t/2, \sqrt{2}t/2, \pi t/4)$. Calculemos la integral del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\mu}$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} &= \int_0^1 \Big[\Big(\tfrac{\sqrt{2}}{2}t\Big)\Big(\tfrac{\pi}{4}t\Big)\Big(\tfrac{\sqrt{2}}{2} - 1\Big) + \Big(1 + \Big(\tfrac{\sqrt{2}}{2} - 1\Big)t\Big)\Big(\tfrac{\pi}{4}t\Big)\Big(\tfrac{\sqrt{2}}{2}\Big) + \Big(1 + \Big(\tfrac{\sqrt{2}}{2} - 1\Big)t\Big)\Big(\tfrac{\sqrt{2}}{2}t\Big)\Big(\tfrac{\pi}{4}\Big)\Big] dt \\ &= \frac{\pi}{8}\Big[\sqrt{2}t^2 - (\sqrt{2} - 1)t^3\Big]_0^1 = \frac{\pi}{8}\end{aligned}$$

Llamamos la atención al hecho de que el valor de la integral resultó el mismo que en el ejemplo anterior. Se integró el mismo campo sobre dos caminos que, siendo distintos, compartían su punto inicial y final. Esta es la situación que se presentó en los ejemplos 5 y 6. Volvemos a plantear la pregunta: ¿una integral de línea de un campo F sobre un camino $\boldsymbol{\lambda}$, depende solamente de los puntos

inicial y final del camino, o depende también de "la forma" de la curva descrita por el camino? La respuesta la daremos en la próxima sección. ■

Ejercicios (Capítulo 7, Sección 3)

En los ejercicios 1–12, calcule las integrales de línea indicadas.

1. $\int_{\lambda} dx$, donde $\lambda: [0, 1] \to \mathbb{R}^2$, $\lambda(t) = (t, 3t)$.

2. $\int_{\lambda} dy$, donde $\lambda: [0, 2] \to \mathbb{R}^2$, $\lambda(t) = (t, 1)$.

3. $\int_{\lambda} x\,dx + y\,dy$, donde $\lambda: [-1, 1] \to \mathbb{R}^2$, $\lambda(t) = (t, t)$.

4. $\int_{\lambda} y\,dx + x\,dy$, donde $\lambda: [-1, 1] \to \mathbb{R}^2$, $\lambda(t) = (-t, t)$.

5. $\int_{\lambda} xy\,dx - y\,dy$, donde $\lambda: [0, 2] \to \mathbb{R}^2$, $\lambda(t) = (t^2, t)$.

6. $\int_{\lambda} (x + y)\,dx + (x - y)\,dy$, donde $\lambda: [0, 2] \to \mathbb{R}^2$, $\lambda(t) = (t, 2t)$.

7. $\int_{\lambda} (x^2 + y^2)\,dx$, donde $\lambda: [0, 1] \to \mathbb{R}^2$, $\lambda(t) = (t, \operatorname{sen}^2 t)$.

8. $\int_{\lambda} (x^2 + y^2)\,dx + (x^2 - y^2)\,dy$, donde $\lambda: [0, 2\pi] \to \mathbb{R}^2$, $\lambda(t) = (\cos t, \operatorname{sen} t)$.

9. $\int_{\lambda} xy\,dx + xz\,dy - yz\,dz$, donde $\lambda: [0, 1] \to \mathbb{R}^3$, $\lambda(t) = (t, t^2, t^3)$.

10. $\int_{\lambda} (x + 2y + z)\,dx + 2y\,dy + (3x - z)\,dz$, donde $\lambda: [0, 2] \to \mathbb{R}^3$, $\lambda(t) = (t + 1, 2t + 1, t)$.

11. $\int_{\lambda} x_1\,dx_1 - x_2\,dx_2 + x_3\,dx_3 - x_4\,dx_4$, donde $\lambda: [0, 2] \to \mathbb{R}^4$, $\lambda(t) = (t, 2t, 3t, 4t)$.

12. $\int_{\lambda} (x_1 + x_2)\,dx_1 - (2x_3 - x_4)\,dx_2 + 2x_3x_4\,dx_3 - (x_4 + x_5)\,dx_4 + x_5\,dx_5$, donde $\lambda: [0, 1] \to \mathbb{R}^5$, $\lambda(t) = (t + 1, t - 1, t + 2, t, 3t)$.

13. Calcule la integral de línea $\int_C 2xy\,dx - y\,dy$, donde C es una curva que une el punto $\mathbf{p} = (0, 0)$ con el punto $\mathbf{q} = (1, 1)$, en cada uno de los casos siguientes:

a. C es un segmento de recta que va de $\mathbf{p}$ a $\mathbf{q}$.

b. C es el arco de la parábola $y = x^2$ que va de $\mathbf{p}$ a $\mathbf{q}$.

c. C es el arco de la parábola $x = y^2$ que va de $\mathbf{p}$ a $\mathbf{q}$.

d. C va de $\mathbf{p}$ al punto $(1, 0)$, y de éste a $\mathbf{q}$ (en línea recta).

e. C va de $\mathbf{p}$ al punto $(0, 1)$, y de éste a $\mathbf{q}$ (en línea recta).

14. Repita el ejercicio anterior con la integral $\int_{\boldsymbol{\lambda}} 6xy\,dx + (3x^2 + 2y)\,dy$.

15. Calcule la integral del ejercicio anterior a lo largo de:

a. el círculo $x^2 + y^2 = r^2$, recorrido en sentido antihorario.

b. el cuadrado $|x| + |y| = 1$, recorrido en sentido horario.

16. Calcule la integral de línea $\int_{\boldsymbol{\lambda}} 3xy\,dx + y^2\,dy$, a lo largo de los caminos de los incisos **a** y **b** del ejercicio anterior.

17. Calcule la integral de línea $\int_{\boldsymbol{\lambda}}(2xy^3 + yz)\,dx + (3x^2y^2 + xz)\,dy + xy\,dz$ donde $\boldsymbol{\lambda}$ es un camino cuyo punto inicial es $\mathbf{p} = (0, 0, 0)$ y cuyo punto final es $\mathbf{q} = (1, 1, 1)$, en cada uno de los casos siguientes:

a. $\boldsymbol{\lambda}$ es un segmento de recta.

b. $\boldsymbol{\lambda}\colon [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (t, t^2, t^3)$.

c. $\boldsymbol{\lambda}\colon [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (t^\alpha, t^\beta, t^\gamma)$, donde $\alpha, \beta, \gamma \in \mathbb{N}$.

18. Calcular la integral de línea del ejercicio anterior a lo largo de la curva cerrada que se obtiene al intersectar las esferas $x^2 + y^2 + z^2 = r^2$, $x^2 + y^2 + (z - r)^2 = r^2$.

19. (Continuación del ejercicio 34 de la sección 12 del capítulo 2: un breve curso "hágalo usted mismo" de funciones de variable compleja. parte (III): integración compleja).

s. Considere la función $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$, $f(z) = u(x, y) + iv(x, y)$, definida en el conjunto abierto U de $\mathbb{C}$. Sea $\boldsymbol{\lambda}\colon [a, b] \to \mathbb{R}^2$ un camino de clase $\mathscr{C}^1$ tal que su imagen esté contenida en U (viendo a éste como un conjunto del plano $\mathbb{R}^2$). Siendo $z = x + iy$, consideremos la "diferencial" de z, de manera formal, como $dz = dx + idy$, y hagamos el producto formal $f(z)\,dz$ como

$$\begin{aligned} f(z)\,dz &= (u(x, y) + iv(x, y))(dx + idy) \\ &= u(x, y)\,dx - v(x, y)\,dy + i(v(x, y)\,dx + u(x, y)\,dy) \end{aligned}$$

La integral de la función $f(z)$ se define a lo largo del camino $\boldsymbol{\lambda}$, denotada por $\int_{\boldsymbol{\lambda}} f(z)\,dz$, como

$$\int_{\boldsymbol{\lambda}} f(z)\,dz = \int_{\boldsymbol{\lambda}} u(x, y)\,dx - v(x, y)\,dy + i\int_{\boldsymbol{\lambda}} v(x, y)\,dx + u(x, y)\,dy$$

(nótese que $\int_{\boldsymbol{\lambda}} f(z)\,dz$ es un número complejo). En términos de las integrales de línea de campos vectoriales que estudiamos en esta sección, podemos escribir

$$\int_{\boldsymbol{\lambda}} f(z)\,dz = \int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} + i\int_{\boldsymbol{\lambda}} G \cdot d\boldsymbol{\lambda}$$

donde $\mathbf{F}, \mathbf{G}\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ son los campos $\mathbf{F}(x, y) = (u(x, y), -v(x, y))$, $\mathbf{G}(x, y) = (v(x, y), u(x, y))$. Si $f, g\colon U \subseteq \mathbb{C} \to \mathbb{C}$ son dos funciones definidas en el abierto U de $\mathbb{C}$, defina su suma $f + g$, y el producto de f por el número real c, de modo natural (como se hace con funciones reales de variable real). Demuestre que

$$\int_{\boldsymbol{\lambda}} (f + cg)(z)\,dz = \int_{\boldsymbol{\lambda}} f(z)\,dz + c\int_{\boldsymbol{\lambda}} g(z)\,dz$$

$$\int_{-\boldsymbol{\lambda}} f(z)\,dz = -\int_{\boldsymbol{\lambda}} f(z)\,dz$$

donde $-\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $(-\boldsymbol{\lambda})(t) = \boldsymbol{\lambda}(a + b - t)$. Si $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, demuestre que

$$\int_{\boldsymbol{\lambda}} f(z)\, dz = \int_{\boldsymbol{\lambda}_1} f(z)\, dz + \int_{\boldsymbol{\lambda}_2} f(z)\, dz$$

t. Calcule las siguientes integrales

t1. $\int_{\boldsymbol{\lambda}} z\, dz$, donde $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$.

t2. $\int_{\boldsymbol{\lambda}} (3z + 2)\, dz$, donde $\boldsymbol{\lambda}: [-2, 2] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t + 2, 3t - 1)$.

t3. $\int_{\boldsymbol{\lambda}} 2z^2\, dz$, donde $\boldsymbol{\lambda}$ es el camino cuya imagen es el triángulo con vértices en $A = (0, 0)$, $B = (1, 3)$, $C = (0, 4)$, recorrido en sentido antihorario.

t4. $\int_{\boldsymbol{\lambda}} 2z^2\, dz$, donde $\boldsymbol{\lambda}$ es el camino cuya imagen es el círculo $x^2 + y^2 = 1$, recorrido en sentido antihorario.

t5. $\int_{\boldsymbol{\lambda}} (z^3 + iz + 2 - i)\, dz$, donde $\boldsymbol{\lambda}$ es el camino cuya imagen es el arco de parábola cúbica $y = x^3$ comprendido entre $x = -1$ y $x = 1$, recorrido del punto $(-1, -1)$ al punto $(1, 1)$.

u. Calcule la integral $\int_{\boldsymbol{\lambda}} z^2\, dz$, a lo largo del camino $\boldsymbol{\lambda}$ que va del punto $\mathbf{p} = (0, 1)$ al punto $\mathbf{q} = (1, 0)$, en cada uno de los casos siguientes:

u1. $\boldsymbol{\lambda}$ es un segmento de recta.

u2. $\boldsymbol{\lambda}$ es un arco de la parábola $y = 1 - x^2$.

u3. $\boldsymbol{\lambda}$ es un arco de la parábola $x = 1 - y^2$.

v. Repita el ejercicio anterior con las integrales

$$\textbf{v1.} \quad \int_{\boldsymbol{\lambda}} \frac{1}{z}\, dz. \qquad \textbf{v2.} \quad \int_{\boldsymbol{\lambda}} \bar{z}\, dz$$

w. Calcule la integral $\int_{\boldsymbol{\lambda}} \frac{1}{z}\, dz$ a lo largo del camino $\boldsymbol{\lambda}$ cuya imagen es:

w1 el círculo $x^2 + y^2 = 1$ recorrido en sentido antihorario.

w2 el cuadrado $|x| + |y| = 1$ recorrido en sentido antihorario.

w3 el cuadrado con vértices en $A = (1, 0)$, $B = (3, 0)$, $C = (3, 2)$, $D = (1, 2)$, con el recorrido $A \to B \to C \to D \to A$.

x. Repita el ejercicio anterior con la integral $\int_{\boldsymbol{\lambda}} \bar{z}\, dz$.
(Este ejercicio continuará en el ejercicio 26 de la próxima sección).

7.4 Independencia del camino, campos conservativos y funciones potenciales

En la sección 2 quedó abierta una pregunta planteada a raíz del ejemplo 6 ahí presentado: dado un campo vectorial $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, ¿hay alguna función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ tal que

grad $f(\mathbf{x}) = \mathbf{F}(\mathbf{x})$, $\mathbf{x} \in U$? Por otra parte, los ejemplos de la sección anterior promovieron que se planteara otra interrogante: al integrar un campo continuo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ sobre un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ de clase $\mathscr{C}^1$, ¿el valor de la integral de línea $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ depende sólo de los extremos $\boldsymbol{\lambda}(a)$ y $\boldsymbol{\lambda}(b)$ del camino $\boldsymbol{\lambda}$, o depende también del mismo camino $\boldsymbol{\lambda}(t)$, $a \leq t \leq b$? En esta sección ahondaremos sobre las cuestiones que estas dos preguntas plantean. Veremos (en el teorema que se presenta a continuación) que estas dos preguntas son en realidad ... ¡la misma pregunta!

Por principio, tenga presente que en caso de que el campo $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ sea el campo gradiente de una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, si el campo $\mathbf{F}$ es de clase $\mathscr{C}^k$, la función f deberá ser de clase $\mathscr{C}^{k+1}$, $k \geq 0$. En efecto, las funciones coordenadas de $\mathbf{F}$ son las derivadas parciales de la función f, de modo que teniendo f derivadas parciales de clase $\mathscr{C}^k$, esta función será de clase $\mathscr{C}^{k+1}$.

Establezcamos el primer resultado importante de esta sección.

Teorema 7.4.1 Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un campo de clase $\mathscr{C}^k$ ($k \geq 0$) definido en el conjunto abierto U de $\mathbb{R}^n$. Las afirmaciones siguientes son equivalentes:

1. $\mathbf{F}$ es el campo gradiente de una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ de clase $\mathscr{C}^{k+1}$.
2. La integral $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ del campo F a lo largo de un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ seccionalmente $\mathscr{C}^1$ (tal que $\boldsymbol{\lambda}([a, b]) \subset U$), depende solamente del punto inicial $\boldsymbol{\lambda}(a)$ y final $\boldsymbol{\lambda}(b)$ del camino $\boldsymbol{\lambda}$.
3. La integral $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ del campo $\mathbf{F}$ a lo largo de un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ cerrado seccionalmente $\mathscr{C}^1$ (de modo que $\boldsymbol{\lambda}([a, b]) \subset U$) es igual a cero.

Demostración. Se probará que $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 2 \Rightarrow 1$.

$1 \Rightarrow 2$ Supongamos entonces que $\mathbf{F} = \operatorname{grad} f$. Escribiendo $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, donde entonces $F_i = \frac{\partial f}{\partial x_i}$, y $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Supongamos por el momento que el camino $\boldsymbol{\lambda}$ es de clase $\mathscr{C}^1$. Se tiene

$$\begin{aligned} \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_a^b \operatorname{grad} f(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\, dt = \int_a^b \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\boldsymbol{\lambda}(t))\lambda_i'(t)\, dt \\ &= \int_a^b \frac{d}{dt} f(\boldsymbol{\lambda}(t))\, dt = \Big[f(\boldsymbol{\lambda}(t)) \Big]_a^b = f(\boldsymbol{\lambda}(b)) - f(\boldsymbol{\lambda}(a)) \end{aligned}$$

de donde se ve que la integral $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ depende solamente de los extremos $\boldsymbol{\lambda}(a)$ y $\boldsymbol{\lambda}(b)$ del camino $\boldsymbol{\lambda}$. Si el camino $\boldsymbol{\lambda}$ es seccionalmente $\mathscr{C}^1$, tomamos una partición $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ de modo que $\boldsymbol{\lambda}$ sea de clase $\mathscr{C}^1$ en cada subintervalo $[t_{j-1}, t_j]$ y repetimos el argumento anterior en cada uno de estos subintervalos. Se obtiene que

$$\begin{aligned} \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \sum_{j=1}^n \int_{\boldsymbol{\lambda}_j} \mathbf{F} \cdot d\boldsymbol{\lambda} = \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \operatorname{grad} f(\lambda_j(t)) \cdot \lambda_j'(t) dt \\ &= \sum_{j=1}^n f(\boldsymbol{\lambda}(t_j)) - f(\boldsymbol{\lambda}(t_{j-1})) = f(\boldsymbol{\lambda}(b)) - f(\boldsymbol{\lambda}(a)) \end{aligned}$$

$2 \Rightarrow 3$ Sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino cerrado de clase $\mathscr{C}^1$. Como observamos en la sección 1, podemos escribir el camino $\boldsymbol{\lambda}$ como $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$. Nótese que los caminos $\boldsymbol{\lambda}_1$ y $-\boldsymbol{\lambda}_2$ tienen los

mismos extremos, de modo que, según **2**, las integrales de línea del campo **F** sobre estos dos caminos valen lo mismo. Se tiene entonces que

$$\begin{aligned}\oint_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda} &= \int_{\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2} F\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_2} \mathbf{F}\cdot d\boldsymbol{\lambda} \\ &= \int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda} - \int_{-\boldsymbol{\lambda}_2} \mathbf{F}\cdot d\boldsymbol{\lambda} = 0\end{aligned}$$

Las modificaciones a este argumento, en caso de que el camino sea seccionalmente $\mathscr{C}^1$, quedan a cargo del lector.

3 $\Rightarrow$ **2** Supongamos que $\boldsymbol{\lambda}_1$ y $\boldsymbol{\lambda}_2$ son dos caminos seccionalmente $\mathscr{C}^1$ que comparten su punto inicial y su punto final. Entonces el camino $\boldsymbol{\mu} = \boldsymbol{\lambda}_1 + (-\boldsymbol{\lambda}_2)$ es un camino cerrado seccionalmente $\mathscr{C}^1$. Se tiene, con **3** que

$$0 = \oint_{\boldsymbol{\mu}} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1+(-\boldsymbol{\lambda}_2)} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda} + \int_{-\boldsymbol{\lambda}_2} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda} - \int_{\boldsymbol{\lambda}_2} \mathbf{F}\cdot d\boldsymbol{\lambda}$$

de donde

$$\int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_2} \mathbf{F}\cdot d\boldsymbol{\lambda}$$

Es decir, la integral $\int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda}$ depende solamente de las extremidades del camino $\boldsymbol{\lambda}_1$.

2 $\Rightarrow$ **1** Esta es la parte más delicada de la demostración. Tendremos que usar algunos hechos técnicos sobre subconjuntos de $\mathbb{R}^n$, los cuales dejaremos para su comentario al final de este teorema. Por lo pronto, vayamos directamente a la idea central de la prueba. Fijemos primeramente un punto $\mathbf{p} \in U$. Para cualquier otro punto $\mathbf{x} \in U$, consideremos el camino $\boldsymbol{\lambda}_{\mathbf{x}}$ cuyo punto inicial esté en $\mathbf{p}$ y su punto final en $\mathbf{x}$, y tal que su imagen quede contenida completamente en U (Atención: ¿quién asegura que este camino existe? Recomendamos al lector que tenga un poco de paciencia: tal camino se da en cada "pedazo" —llamado "componente conexa"— de U). Establezcamos la función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida como

$$f(\mathbf{x}) = \int_{\boldsymbol{\lambda}_{\mathbf{x}}} \mathbf{F}\cdot d\boldsymbol{\lambda}$$

Nótese que esta función *está bien definida,* pues suponemos que el valor de la integral (es decir, el valor de $f(\mathbf{x})$) depende solamente de los extremos del camino $\boldsymbol{\lambda}_{\mathbf{x}}$ que son el punto $\mathbf{p}$ que está fijo, y el punto $\mathbf{x}$. La idea será ahora demostrar que el campo $\mathbf{F}$ es el campo gradiente de esta función f. Para esto debemos probar que si $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, entonces $F_i = \frac{\partial f}{\partial x_i}$, $i = 1, 2, \ldots, n$, en cada punto $\mathbf{x} \in U$. Calculemos entonces la i-ésima derivada parcial de la función f en un punto arbitrario $\mathbf{x} \in U$. Se tiene

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h}$$

donde $\mathbf{e}_i$ es el i-ésimo vector de la base canónica de $\mathbb{R}^n$. Siendo U un conjunto abierto de $\mathbb{R}^n$, para el punto $\mathbf{x} \in U$ existe una bola $B_{\mathbf{x}}$ con centro en $\mathbf{x}$ y radio $\epsilon > 0$ tal que $B_{\mathbf{x}} \subset U$. Tomando entonces h suficientemente pequeño, podemos asegurar que $\mathbf{x} + h\mathbf{e}_i$ es un elemento de la bola $B_{\mathbf{x}}$, y por lo tanto de U. Obsérvese que

$$f(\mathbf{x} + h\mathbf{e}_i) = \int_{\boldsymbol{\lambda}_{\mathbf{x}+h\mathbf{e}_i}} \mathbf{F}\cdot d\boldsymbol{\lambda}$$

donde $\boldsymbol{\lambda}_{\mathbf{x}+h\mathbf{e}_i}$ es un camino que une al punto $\mathbf{p}$ con el punto $\mathbf{x}+h\mathbf{e}_i$. Este camino se puede considerar como la suma del camino $\boldsymbol{\lambda}_{\mathbf{x}}$ que une al punto $\mathbf{p}$ con $\mathbf{x}$, y un camino $\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}$ que una al punto x con el punto $\mathbf{x}+h\mathbf{e}_i$. Entonces, poniendo $\boldsymbol{\lambda}_{\mathbf{x}+h\mathbf{e}_i} = \boldsymbol{\lambda}_{\mathbf{x}} + \boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}$, nos queda que

$$
\begin{aligned}
f(\mathbf{x}+h\mathbf{e}_i) - f(\mathbf{x}) &= \int_{\boldsymbol{\lambda}_{\mathbf{x}+h\mathbf{e}_i}} \mathbf{F}\cdot d\boldsymbol{\lambda} - \int_{\boldsymbol{\lambda}_{\mathbf{x}}} \mathbf{F}\cdot d\boldsymbol{\lambda} \\
&= \int_{\boldsymbol{\lambda}_{\mathbf{x}}+\boldsymbol{\mu}_{\mathbf{x}+h\mathbf{e}_i}} \mathbf{F}\cdot d\boldsymbol{\lambda} - \int_{\boldsymbol{\lambda}_{\mathbf{x}}} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_{\mathbf{x}}} \mathbf{F}\cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}} \mathbf{F}\cdot d\boldsymbol{\mu} - \int_{\boldsymbol{\lambda}_{\mathbf{x}}} \mathbf{F}\cdot d\boldsymbol{\lambda} \\
&= \int_{\boldsymbol{\mu}_{\mathbf{x}+h\mathbf{e}_i}} \mathbf{F}\cdot d\boldsymbol{\lambda}
\end{aligned}
$$

donde, como decíamos, $\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}$ es un camino que une el punto $\mathbf{x}$ con el punto $\mathbf{x}+h\mathbf{e}_i$. Por las propiedades que tienen las bolas en $\mathbb{R}^n$ (si dos puntos pertenecen a una bola, el segmento de recta que une a estos puntos sigue perteneciendo a la bola; propiedad conocida como *convexidad*), podemos tomar el camino $\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}$ como $\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}\colon [0,h] \to \mathbb{R}^n$, $\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}(t) = \mathbf{x}+t\mathbf{e}_i$. Entonces

$$
f(\mathbf{x}+h\mathbf{e}_i) - f(\mathbf{x}) = \int_{\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}} \mathbf{F}\cdot d\boldsymbol{\mu} = \int_0^h \mathbf{F}(\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}(t))\cdot \boldsymbol{\mu}'_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}(t)\,dt
$$

Nótese que $\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}(t) = \mathbf{e}_i$, de modo que el integrando que aparece en la última integral de la expresión anterior se convierte en $F_i(\boldsymbol{\mu}_{\mathbf{x},\mathbf{x}+h\mathbf{e}_i}(t)) = F_i(\mathbf{x}+t\mathbf{e}_i)$, por lo que

$$
f(\mathbf{x}+h\mathbf{e}_i) - f(\mathbf{x}) = \int_0^h F_i(\mathbf{x}+t\mathbf{e}_i)\,dt
$$

y entonces

$$
\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{e}_i) - f(\mathbf{x})}{h} = \lim_{h\to 0} \frac{1}{h}\int_0^h F_i(\mathbf{x}+t\mathbf{e}_i)\,dt
$$

Por otra parte, si consideramos la función $\phi\colon \mathbb{R} \to \mathbb{R}$,

$$
\phi(s) = \int_0^s F_i(\mathbf{x}+t\mathbf{e}_i)\,dt
$$

y calculamos su derivada en $s=0$, nos queda

$$
\begin{aligned}
\phi'(0) &= \lim_{h\to 0} \frac{\phi(h)-\phi(0)}{h} = \lim_{h\to 0} \frac{1}{h}\phi(h) \\
&= \lim_{h\to 0} \frac{1}{h}\int_0^h F_i(\mathbf{x}+t\mathbf{e}_i)\,dt = \frac{\partial f}{\partial x_i}(\mathbf{x})
\end{aligned}
$$

Pero segun el Teorema Fundamental del Calculo se tiene

$$
\begin{aligned}
\phi'(0) &= \Big(\phi'(s)\Big)_{s=0} = \left(\frac{d}{ds}\int_0^s F_i(\mathbf{x}+t\mathbf{e}_i)\,dt\right)_{s=0} \\
&= \Big(F_i(x+se_i)\Big)_{s=0} = F_i(\mathbf{x})
\end{aligned}
$$

Es decir que $\frac{\partial f}{\partial x_i}(\mathbf{x}) = F_i(\mathbf{x})$, como queríamos probar. Q.E.D.

En la demostración de **2** $\Rightarrow$ **1** en el teorema anterior, aparecen varias cuestiones delicadas (desde la perspectiva del rigor matemático) que no se aclararon, por ejemplo la existencia del camino $\boldsymbol{\lambda}_{\mathbf{x}}$ que se usó en el argumento. Estos conflictos se pueden explicar con un poco de topología de conjuntos en el espacio $\mathbb{R}^n$. Al lector con mentalidad más matemática, que esté ávido de ver cómo resuelve la matemática este tipo de problemas, le recomendamos estudiar la próxima sección. Los campos con las características mencionadas en el teorema anterior reciben un nombre especial.

Definición. Al campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, de clase $\mathscr{C}^k$, $k \geq 0$, definido en el abierto U de $\mathbb{R}^n$, que cumpla con alguna de (y por lo tanto con todas) las condiciones del teorema anterior, se le llama *campo conservativo,* y a la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ de clase $\mathscr{C}^{k+1}$ tal que $\mathbf{F} = \operatorname{grad} f$, se le llama *función potencial.* ■

Así pues, en el teorema 7.4.1 se ha probado que un campo conservativo se caracteriza por la existencia de una función potencial, o bien, de modo equivalente, por el hecho de que la integral del campo a lo largo de un camino dependa sólo de los extremos de este último, o bien, por que la integral del campo sobre cualquier camino cerrado sea cero.

Corolario de la demostración del teorema anterior). Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathscr{C}^k$, $k \geq 1$, y sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino seccionalmente $\mathscr{C}^1$. Entonces

$$\int_{\boldsymbol{\lambda}} \operatorname{grad} f \cdot d\boldsymbol{\lambda} = f(\boldsymbol{\lambda}(b)) - f(\boldsymbol{\lambda}(a))$$ ■

La demostración de este corolario se efectuó en la demostración de **1** $\Rightarrow$ **2** en el teorema 7.4.1. Nótese lo que este resultado establece en el caso $n = 1$. Se tiene una función $f: U \subseteq \mathbb{R} \to \mathbb{R}$ (real de una variable real) de clase $\mathscr{C}^k$, $k \geq 1$, y un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}$ seccionalmente $\mathscr{C}^1$. Siendo $\boldsymbol{\lambda}$ una función continua definida en el intervalo compacto $[a, b]$, (se puede demostrar que) su imagen $\boldsymbol{\lambda}([a, b]) \subset \mathbb{R}$ es también un intervalo compacto, digamos $[\alpha, \beta]$, donde $\alpha = \boldsymbol{\lambda}(a)$, $\beta = \boldsymbol{\lambda}(b)$. Por otra parte, es claro que en este caso $\operatorname{grad} f(\mathbf{x}) = f'(\mathbf{x})$. Entonces, escribiendo $\mathbf{x} = \boldsymbol{\lambda}(t)$, la fórmula del corolario anterior se ve como

$$\int_{\boldsymbol{\lambda}} \operatorname{grad} f \cdot d\boldsymbol{\lambda} = \int_{\alpha}^{\beta} f'(x)\, dx = f(\beta) - f(\alpha)$$

que no es más que el Teorema Fundamental del Cálculo que ya conocemos. Así pues, la fórmula del corolario anterior es una generalización de este teorema. En tal fórmula identificamos al gradiente de f (que es un campo) como "la derivada" de la función f, siendo ahora la región de integración un camino con extremos en $\boldsymbol{\lambda}(a)$ y $\boldsymbol{\lambda}(b)$. Lo que dice tal fórmula es que "la integral de la derivada de f sobre una región con punto inicial $\boldsymbol{\lambda}(a)$ y punto final $\boldsymbol{\lambda}(b)$ es el valor de la función f en el punto final $\boldsymbol{\lambda}(b)$ menos el valor de f en el punto inicial $\boldsymbol{\lambda}(a)$". Esto es exactamente lo que nos dice el Teorema Fundamental del Cálculo para funciones reales de variable real. Seguiremos usando este nombre para designar al resultado del corolario anterior.

Antes de iniciar el estudio de algunos ejemplos, téngase presente que las funciones potenciales $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ de un campo conservativo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ no son únicas (como tampoco lo son las primitivas de las funciones reales de variable real), pero, como es de esperarse, si f y g son dos funciones potenciales del campo $\mathbf{F}$ definidas en el conjunto U abierto y conexo (de modo que U esté constituido por una sola pieza, como ocurre con los intervalos de $\mathbb{R}$; los detalles técnicos a este

respecto se discutirán en la próxima sección), éstas difieren por una constante. Dejamos al lector la verificación este hecho.

Ejemplo 1. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y^2, 2xy)$. Este es un campo conservativo, pues la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \frac{1}{2}x^2 + y^2 x$ es una función potencial de $\mathbf{F}$. En efecto, se tiene que

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \left(x + y^2, 2xy\right) = \mathbf{F}(x, y)$$

De hecho, en los ejemplos 5 y 6 de la sección anterior se obtuvo la integral de este campo a lo largo de los caminos $\boldsymbol{\lambda}, \boldsymbol{\nu}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$, $\boldsymbol{\nu}(t) = (t, t^3)$, los cuales comparten sus extremos. Se obtuvo que

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\nu}} \mathbf{F}\, d\boldsymbol{\nu} = \frac{3}{2}$$

tal como asegura el teorema anterior que debe ocurrir (la integral solamente depende del punto inicial $\boldsymbol{\lambda}(0) = \boldsymbol{\nu}(0) = (0, 0)$ y del punto final $\boldsymbol{\lambda}(1) = \boldsymbol{\nu}(1) = (1, 1)$). Más aún, usando el corolario del teorema anterior con la función potencial $f(x, y) = \frac{1}{2}x^2 + y^2 x$, $\boldsymbol{\lambda}(a) = \boldsymbol{\lambda}(0) = (0, 0)$, $\boldsymbol{\lambda}(b) = \boldsymbol{\lambda}(1) = (1, 1)$, tenemos

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_{\boldsymbol{\lambda}} \operatorname{grad} f \cdot d\boldsymbol{\lambda} = \left(\frac{1}{2}x^2 + y^2 x\right)_{\substack{x=1\\y=1}} - \left(\frac{1}{2}x^2 + y^2 x\right)_{\substack{x=0\\y=0}} \\ &= \frac{1}{2} + 1 = \frac{3}{2}\end{aligned}$$

■

Ejemplo 2. El campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y, y)$ no es conservativo. Esto se puede concluir fácilmente de los ejemplos 1 y 4 de la sección anterior. En efecto, en el ejemplo 1 se obtuvo la integral de $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$, llegando a

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \frac{4}{3}$$

y en el ejemplo 4, se obtuvo la integral de $\mathbf{F}$ a los largo del camino $\boldsymbol{\nu}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\nu}(t) = (t, t^3)$,

$$\int_{\boldsymbol{\nu}} \mathbf{F} d\boldsymbol{\nu} = \frac{5}{4}$$

Los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\nu}$ tienen los mismos extremos, pero la integral de $\mathbf{F}$ a lo largo de cada uno de ellos es diferente. Por el teorema 7.4.1, concluimos entonces que el campo $\mathbf{F}$ no es conservativo. ■

Ejemplo 3. El campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (2x + y, -x + 4y)$ no es conservativo. En efecto, en el ejemplo 8 de la sección anterior se obtuvo la integral de $\mathbf{F}$ a lo largo del círculo $x^2 + y^2 = 1$, recorrido (una vez) en sentido antihorario, comenzando en $(1, 0)$, el cual se puede ver como (la imagen de) un camino cerrado simple. Se obtuvo que

$$\oint_{x^2+y^2=1} F_1(x, y)\, dx + F_2(x, y)\, dy = -2\pi \neq 0$$

Por el teorema 7.4.1 (inciso *c*), concluimos que el campo no es conservativo. ■

Ejemplo 4. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + 4y, ax + y)$. Este campo es conservativo si y sólo si $a = 4$. Este es un hecho del cual nos podremos convencer más adelante (en el ejemplo 12). Por lo pronto, nótese que, en tal caso, una función potencial del campo $\mathbf{F}(x, y) = (x + 4y, 4x + y)$ es $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \frac{1}{2}(x^2 + y^2) + 4xy$, pues

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (x + 4y, 4x + y) = F(x, y)$$

Además, en el ejemplo 9 de la sección anterior se obtuvo la integral de $\mathbf{F}$ a lo largo del camino cerrado simple seccionalmente $\mathscr{C}^1$ cuya traza es la curva $C = \{(x, y) | |x| + |y| = 1\}$ obteniéndose el valor $\frac{1}{2}a - 2$. Se observa que si $a = 4$, la integral es cero, como asegura el teorema 7.4.1 que tiene que ocurrir. Del modo semejante, en el ejemplo 10 de la sección anterior, se calculó la integral de $\mathbf{F}$ a lo largo del círculo $x^2 + y^2 = r^2$ recorrido en sentido antihorario, obteniéndose el valor $\pi r^2(a - 4)$, valor que es cero con $a = 4$. ■

Ejemplo 5. El campo $\mathbf{F}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$, dado por

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

no es conservativo. En efecto, en el ejemplo 11 de la sección anterior se calculó la integral de $\mathbf{F}$ a lo largo del camino que describe un arco del círculo $x^2 + y^2 = r^2$ correspondiente al ángulo central $k\pi$, donde $0 < k \leq 2$ recorrido en sentido antihorario, obteniéndose el valor de la integral igual a $k\pi$. En particular, si $k = 2$, se tiene un camino cerrado simple, y en tal caso la integral de $\mathbf{F}$ a lo largo de este camino sería $2\pi \neq 0$. ■

Ejemplo 6. Consideremos el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (yz, xz, xy)$. Este es un campo conservativo, pues la función $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = xyz$ es una función potencial de F. En efecto

$$\operatorname{grad} f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = (yz, xz, xy) = \mathbf{F}(x, y, z)$$

En los ejemplos 12 y 13 de la sección anterior se obtuvo la integral de $\mathbf{F}$ a lo largo de dos caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\nu}$ que compartían su punto inicial $\boldsymbol{\lambda}(0) = \boldsymbol{\nu}(0) = (1, 0, 0)$ y su punto final $\boldsymbol{\lambda}(\pi/4) = \boldsymbol{\nu}(\pi/4) = (\sqrt{2}/2, \sqrt{2}/2, \pi/4)$, dando como resultado en ambos casos $\pi/8$. De nuevo, nótese que, conociendo la función potencial de $\mathbf{F}$, el valor de la integral se puede calcular como (según el corolario del teorema 7.4.1)

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = f(\boldsymbol{\lambda}(b)) - f(\boldsymbol{\lambda}(a)) = (xyz)_{\substack{x=\sqrt{2}/2 \\ y=\sqrt{2}/2 \\ z=\pi/4}} - (xyz)_{\substack{x=1 \\ y=0 \\ z=0}} = \frac{\pi}{8}$$ ■

Aunque el teorema 7.4.1 nos presenta la equivalencia entre varias condiciones para que un campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ sea conservativo, éstas no son muy prácticas para decidir cuándo sí *es conservativo* el campo. Por ejemplo, no podemos verificar en la práctica que la integral del campo $\mathbf{F}$ a lo largo de *todo* camino cerrado simple sea cero. Ciertamente tal teorema nos ayuda a decidir cuándo un campo $\mathbf{F}$ *no es conservativo*: basta con hallar un camino cerrado simple sobre el que la integral de F no sea

cero, o bien, dos caminos con los mismos extremos sobre los que las integrales de F tengan distintos valores.

Presentaremos ahora otro tipo de condiciones que debe cumplir un campo $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ de clase $\mathscr{C}^k$, con $k \geq 1$, para que sea conservativo. Consideremos por el momento el caso $n = 2$. Tenemos entonces el campo $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, de clase $\mathscr{C}^k$, $k \geq 1$. Escribamos $\mathbf{F}(x, y) = (M(x, y), N(x, y))$, donde entonces las funciones $M, N: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son de clase $\mathscr{C}^k$, $k \geq 1$. Supongamos que $\mathbf{F}$ es conservativo. Entonces hay una función potencial $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, de clase $\mathscr{C}^{k+1}$ de modo que

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right) = \big(M(x, y), N(x, y)\big) = F(x, y)$$

es decir, que

$$\frac{\partial f}{\partial x}(x, y) = M(x, y), \qquad \frac{\partial f}{\partial y}(x, y) = N(x, y)$$

para toda (x, y) en U. Derivando la primera de estas expresiones respecto de y y la segunda respecto de x obtenemos

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x, y)\right) = \frac{\partial}{\partial y}\big(M(x, y)\big) = \frac{\partial M}{\partial y}$$

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x, y)\right) = \frac{\partial}{\partial x}\big(N(x, y)\big) = \frac{\partial N}{\partial x}$$

Ahora bien, puesto que la función f es al menos de clase $\mathscr{C}^2$, el teorema de Schwarz (Teorema 2.12.1) nos asegura que las derivadas parciales cruzadas de f son iguales. Entonces, siendo el campo $\mathbf{F} = (M, N)$ conservativo, se debe tener que

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

en todo punto (x, y) de U.

Ejemplo 7. En el ejemplo 1 se vio que el campo (de clase $\mathscr{C}^\infty$) $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x+y^2, 2xy)$ es conservativo. En este caso se tiene $M(x, y) = x + y^2$, $N(x, y) = 2xy$, y

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(x + y^2) = 2y = \frac{\partial}{\partial x}(2xy) = \frac{\partial N}{\partial x} \qquad \blacksquare$$

Ejemplo 8. En el ejemplo 2 se vio que el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (x + y, y)$ no es conservativo. Esto se puede concluir también del hecho de que

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(x + y) = 1 \neq 0 = \frac{\partial}{\partial x}(y) = \frac{\partial N}{\partial x}$$

De modo semejante, el campo del ejemplo 3, $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (2x + y, -x + 4y)$ no es conservativo, pues si lo fuera tendría que cumplir la condición $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$, pero

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(2x + y) = 1 \neq -1 = \frac{\partial}{\partial x}(-x + 4y) = \frac{\partial N}{\partial x} \qquad \blacksquare$$

Consideremos ahora el caso general. Si $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ es un campo de clase $\mathscr{C}^k$, $k \geq 1$, conservativo, entonces hay una función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ de clase $\mathscr{C}^{k+1}$ que grad $f(\mathbf{x}) = \mathbf{F}(\mathbf{x})$, es decir, que

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = F_i(\mathbf{x}), \qquad \mathbf{x} \in U$$

$i = 1, 2, \ldots, n$, donde $\mathbf{F} = (F_1, F_2, \ldots, F_n)$. Derivando esta expresión respecto de x_j, $i \neq j$, obtenemos

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) = \frac{\partial F_i}{\partial x_j}(\mathbf{x})$$

Intercambiando los papeles de i y de j obtenemos que se debe cumplir también que

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

Siendo f una función de clase al menos $\mathscr{C}^2$, sus parciales cruzadas deben ser iguales. Es decir, se debe tener

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

Hemos probado entonces el resultado siguiente

Teorema 7.4.2 (Condiciones necesarias para que un campo sea conservativo) Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ un campo de clase $\mathscr{C}^k$ $k \geq 1$, definido en el conjunto abierto U de $\mathbb{R}^n$. Si F es conservativo entonces

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

para $x \in U$, $1 \leq i < j \leq n$. ■

Por ejemplo, para que el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ de clase $\mathscr{C}^k$, $k \geq 1$, $\mathbf{F}(x, y, z) = (M(x, y, z), N(x, y, z), P(x, y, z))$ sea conservativo, es necesario que se cumplan las 3 condiciones

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}, \quad \frac{\partial M}{\partial z} = \frac{\partial P}{\partial x}, \quad \frac{\partial N}{\partial z} = \frac{\partial P}{\partial y}$$

para todo $(x, y, z) \in U$.

Ejemplo 9. Consideremos el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (yz, xz, xy)$ del ejemplo 6. Se vio que este campo es conservativo, pues se exhibió una función potencial para él ($f(x, y, z) = xyz$). Nótese que en este caso $M(x, y, z) = yz$, $N(x, y, z) = xz$, $P(x, y, z) = xy$, y como lo asegura el teorema 7.4.2 se cumple que

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(yz) = z = \frac{\partial}{\partial x}(xz) = \frac{\partial N}{\partial x}$$
$$\frac{\partial M}{\partial z} = \frac{\partial}{\partial z}(yz) = y = \frac{\partial}{\partial x}(xy) = \frac{\partial P}{\partial x}$$
$$\frac{\partial N}{\partial z} = \frac{\partial}{\partial z}(xz) = x = \frac{\partial}{\partial y}(xy) = \frac{\partial P}{\partial y}$$

■

La pregunta natural que se plantea al ver el resultado del teorema 7.4.2, es si las condiciones necesarias ahí establecidas para que el campo **F** sea conservativo son también suficientes. Es decir, si el campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ de clase $\mathscr{C}^k$, $k \geq 1$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, es tal que

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

para todo $x \in U$, $1 \leq i < j \leq n$, ¿podemos concluir que este campo es conservativo? La respuesta a esta pregunta es: NO. Es decir, las condiciones establecidas en el teorema 7.4.2 para que **F** sea conservativo son condiciones *necesarias pero no suficientes*. Para convencernos de esto, tomemos el campo $\mathbf{F}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right)$$

Este es un campo de clase $\mathscr{C}^\infty$ en el abierto $U = \mathbb{R}^2 - \{(0, 0)\}$. Llamando como antes $M(x, y)$ y $N(x, y)$ a las funciones coordenadas de **F**, vemos que

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}\left(\frac{-y}{x^2 + y^2}\right) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) = \frac{\partial N}{\partial x}$$

Es decir, se cumple la condición de igualdad entre las derivadas parciales establecida en el teorema 7.4.2, pero, en el ejemplo 5 de esta sección, se vio que este campo no es conservativo, pues la integral de **F** a lo largo del camino cerrado simple $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2 - \{(0, 0)\}$, $\boldsymbol{\lambda}(t) = (r \cos t, r \operatorname{sen} t)$ es igual a $2\pi \neq 0$, y entonces, por la parte **c** del teorema 7.4.1, se concluye que **F** no puede ser conservativo.

Un hecho sobre el que llamamos la atención es que la propiedad del campo **F** de ser conservativo, es una propiedad *global*: se pide que haya una función f definida donde está definido **F**, y que en todo U se tenga que **F** es el campo gradiente de f. Por otra parte, la propiedad establecida en el teorema 7.4.2, que está expresada en términos de las derivadas parciales de las funciones coordenadas de **F**, es una propiedad *local*: tales derivadas parciales establecen un comportamiento determinado del campo **F** en los alrededores del punto en que ocurre la igualdad de las derivadas parciales. No es extraño pues que, en principio, estas dos propiedades *no sean equivalentes*. Lo que sí podemos esperar que acontezca, en base a la observación hecha en este párrafo, es que la propiedad establecida en el teorema 7.4.2 garantice *localmente* que el campo **F** es conservativo.

Se dice que el campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, de clase $\mathscr{C}^k$, $k \geq 0$, es *localmente conservativo,* si para cada punto $\mathbf{p} \in U$ se da una bola $B_\mathbf{p} = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x} - \mathbf{p}\| < \epsilon\}$ en la que el campo es conservativo, es decir, en la que existe una función $f: B_\mathbf{p} \to \mathbb{R}$, de clase $\mathscr{C}^{k+1}$, tal que grad $f(\mathbf{x}) = \mathbf{F}(\mathbf{x})$, $\mathbf{x} \in B_\mathbf{p}$. Se tiene entonces el resultado siguiente

Teorema 7.4.3 Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ un campo de clase $\mathscr{C}^k$, $k \geq 1$, definido en el abierto U de $\mathbb{R}^n$. Suponga que

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

para todo $x \in U$, $1 \leq i < j \leq n$. Entonces el campo **F** es localmente conservativo.

Demostración. Por sencillez, sólo haremos la demostración en el caso $n = 2$. El caso general se prueba de la misma manera. Escribiremos entonces el campo **F** como $\mathbf{F} = (M, N)$ donde $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Sea **p** un punto del abierto U donde se defina **F**. Sin pérdida de generalidad, podemos suponer que **p** es el origen (el caso general se obtiene haciendo una traslación del origen). Siendo U abierto, hay una bola $B_{\mathbf{p}} = \{(x, y) \in \mathbb{R}^2 | \|(x, y)\| < \epsilon\}$ con centro en $\mathbf{p} = (0, 0)$ que está contenida en U. Considere el camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = t(x, y)$. Nótese que este camino conecta con una recta el origen **p** con el punto (x, y). Obsérvese también que si $(x, y) \in B_{\mathbf{p}}$, entonces la imagen de $\boldsymbol{\lambda}$ está contenida en la bola $B_{\mathbf{p}}$. Definamos entonces la función $f: B_{\mathbf{p}} \to \mathbb{R}$ como

$$f(x, y) = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

Demostraremos que grad $f(x, y) = \mathbf{F}(x, y)$, $(x, y) \in B_{\mathbf{p}}$. Se tiene

$$\begin{aligned}
\frac{\partial f}{\partial x}(x, y) &= \frac{\partial}{\partial x} \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \frac{\partial}{\partial x} \int_0^1 \big(M(tx, ty)x + N(tx, ty)y\big) dt \\
&= \int_0^1 \frac{\partial}{\partial x} \big(M(tx, ty)x + N(tx, ty)y\big) dt \\
&= \int_0^1 \left(x\frac{\partial}{\partial x} M(tx, ty) + M(tx, ty) + y\frac{\partial}{\partial x} N(tx, ty) \right) dt \\
&= \int_0^1 \left(xt\frac{\partial M}{\partial x}(tx, ty) + M(tx, ty) + yt\frac{\partial N}{\partial x}(tx, ty) \right) dt \\
&= \int_0^1 \left(xt\frac{\partial M}{\partial x}(tx, ty) + M(tx, ty) + yt\frac{\partial M}{\partial y}(tx, ty) \right) dt \\
&= \int_0^1 \left(t(x\frac{\partial M}{\partial x}(tx, ty) + y\frac{\partial M}{\partial y}(tx, ty)) + M(tx, ty) \right) dt \\
&= \int_0^1 \left(t\frac{d}{dt}(M(tx, ty)) + M(tx, ty) \right) dt \\
&= \int_0^1 \left(\frac{d}{dt}(tM(tx, ty)) \right) dt = \Big[tM(tx, ty) \Big]_0^1 = M(x, y)
\end{aligned}$$

Es decir, hemos probado que

$$\frac{\partial f}{\partial x}(x, y) = M(x, y), \quad (x, y) \in B_{\mathbf{p}}$$

De manera análoga se prueba que

$$\frac{\partial f}{\partial y}(x, y) = N(x, y), \quad (x, y) \in B_{\mathbf{p}}$$

y así

$$\text{grad } f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right) = \big(M(x, y), N(x, y)\big) = F(x, y)$$

para $(x, y) \in B_{\mathbf{p}}$. Es decir, f es una función potencial de $\mathbf{F}$ en $B_{\mathbf{p}}$, y por ende $\mathbf{F}$ es un campo conservativo en la bola $B_{\mathbf{p}}$. Q.E.D.

Algunas observaciones importantes acerca de la demostración del teorema anterior. Nótese primeramente que el éxito de la demostración está basado en que *se exhibió* (... una manera de probar la existencia en matemáticas) una función potencial para el campo **F**, pues con esto probamos que **F** es conservativo. La función potencial f exhibida en la demostración es una función que en el punto (x, y) se definió como la integral de línea del campo **F** a lo largo del camino $\boldsymbol{\lambda}$ cuya imagen es la recta que une al punto **p** (que es fijo) con el punto (x, y). Para lograr esto, necesitamos que la imagen del camino $\boldsymbol{\lambda}$ *esté contenida en el abierto $B_{\mathbf{p}}$ donde definimos la función potencial*. La bola $B_{\mathbf{p}}$ tiene justamente la propiedad de que habiendo dos puntos **p** y **q** de ella, la recta que los une queda completamente dentro de la bola. Por lo tanto, la demostración anterior se podría repetir si el conjunto abierto U en donde está definido el campo **F** tiene esta misma propiedad. En tal caso, la función potencial f se podría definir *globalmente* en todo el abierto U. Estas reflexiones nos hacen sentir que "la forma" del abierto U juega un papel muy importante en el hecho de que las condiciones de igualdad entre las derivadas parciales de las funciones componentes del campo **F** establecidas en el teorema 7.4.2, sean condiciones *suficientes* para que el campo sea conservativo. Por ejemplo, si el campo **F** está definido en una bola B, tales condiciones garantizan que es conservativo. También, si **F** está definido en todo el espacio $\mathbb{R}^n$, se tiene la suficiencia en el teorema 7.4.2. La demostración de estos hechos sería exactamente la misma que se hizo en el teorema 7.4.3, definiendo una función potencial de **F**, *en el mismo abierto donde* **F** *está definida*, como una integral de línea. Más en general, si el abierto U donde está definido el campo **F** tiene la propiedad de que dados dos puntos cualesquiera de él, la recta que los une queda completamente contenida en él mismo, la misma idea de la demostración del teorema 7.4.3 probaría que el campo **F** (con las condiciones ahí establecidas) es conservativo. Este tipo de conjuntos U de $\mathbb{R}^n$ reciben un nombre especial.

Definición. (Conjunto convexo). Se dice que el conjunto $U \subseteq \mathbb{R}^n$ es un *conjunto convexo*, si dados dos puntos $\mathbf{p}, \mathbf{q} \in U$, el segmento de recta que los une $[\mathbf{p}, \mathbf{q}] = \{\mathbf{x} \in \mathbb{R}^n | x = t\mathbf{q} + (1 - t)\mathbf{p}, 0 \leq t \leq 1\}$ queda contenido en U (es decir, $[\mathbf{p}, \mathbf{q}] \subset U$). ■

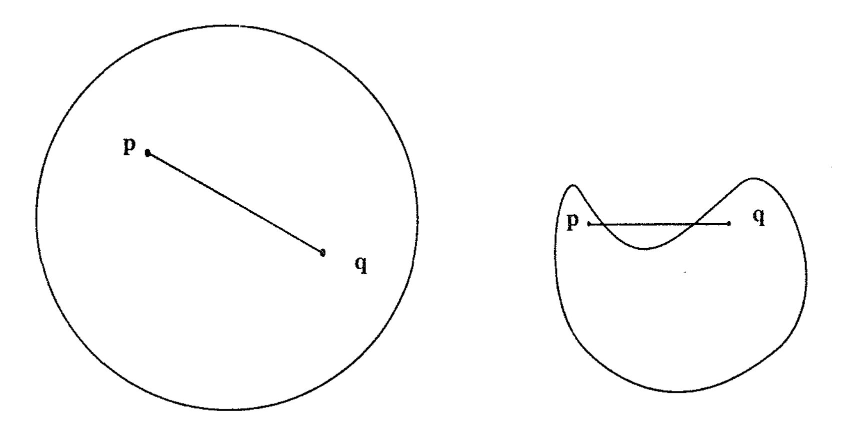


Figura 1. Un conjunto convexo y un conjunto no convexo

Así entonces, las bolas abiertas de $\mathbb{R}^n$ y todo el espacio $\mathbb{R}^n$ son ejemplos de conjuntos convexos. Se tiene entonces el siguiente resultado.

Teorema 7.4.4 Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ un campo de clase $\mathscr{C}^k$, $k \geq 1$, definido en el abierto convexo U de $\mathbb{R}^n$. Una condición necesaria y suficiente para que el campo $\mathbf{F}$ sea conservativo es que

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

para todo $x \in U$, $1 \leq i < j \leq n$.

Demostración. Que la condición es necesaria, ya quedó establecido en el teorema 7.4.2. Veamos que es suficiente. Consideramos el caso $n = 2$. Sea $\mathbf{p}$ un punto arbitrario pero fijo de U. Defina $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ como

$$f(x, y) = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

donde $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = t(x, y) + (1 - t)\mathbf{p}$. Nótese que $\boldsymbol{\lambda}([0, 1]) = [\mathbf{p}, (x, y)]$ = recta que une a $\mathbf{p}$ con (x, y), y que, por ser U convexo, $\boldsymbol{\lambda}([0, 1]) \subset U$, por lo que hace perfecto sentido la definición de f. De la misma manera como se hizo en el teorema 7.4.3, se demuestra que

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right) = \big(M(x, y), N(x, y)\big) = F(x, y)$$

Es decir que f es una función potencial del campo $\mathbf{F}$, y entonces $\mathbf{F}$ es conservativo. ■

Ejemplo 10. El conjunto abierto $U = \mathbb{R}^2 - \{(0, 0)\}$ no es un conjunto convexo. De hecho, tome los puntos $\mathbf{p}$ y $\mathbf{q} \doteq -\mathbf{p}$ en U y observe que la recta que conecta a $\mathbf{p}$ con $\mathbf{q}$ es $[\mathbf{p}, \mathbf{q}] = t\mathbf{q} + (1 - t)\mathbf{p} = -t\mathbf{p} + (1 - t)\mathbf{p} = (1 - 2t)\mathbf{p}$, con $0 \leq t \leq 1$. Para $t = 1/2$, el punto de la recta $[\mathbf{p}, \mathbf{q}]$ es el origen de coordenadas *que no pertenece a* U. Es decir, no se cumple la propiedad $[\mathbf{p}, \mathbf{q}] \subset U$. En este caso entonces el teorema anterior *no garantiza* que las condiciones de igualdad entre las derivadas parciales de un campo definido en U sean suficientes para asegurar que el campo es conservativo. De hecho, tenemos el ejemplo concreto del campo $\mathbf{F}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$, dado por

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right)$$

en el que hemos verificado que se cumplen las condiciones del teorema 7.4.2, pero sabemos que NO es conservativo (ver ejemplo 5). ■

El teorema 7.4.4 nos dice cómo podemos obtener funciones potenciales de campos conservativos. Veamos algunos ejemplos.

Ejemplo 11. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y^2, 2xy)$. Se tiene

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(x + y^2) = 2y = \frac{\partial}{\partial x}(2xy) = \frac{\partial N}{\partial x}$$

Al quedar definido $\mathbf{F}$ en el conjunto abierto convexo $\mathbb{R}^2$, concluimos que el campo es conservativo. Por otra parte, en el ejemplo 1 se verificó que la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \frac{1}{2}x^2 + y^2x$ es una función potencial de $\mathbf{F}$. Veamos cómo obtener esta función. Atendiendo a la demostración del teorema anterior, fijemos el punto $\mathbf{p} = (0, 0)$ y definamos $f: \mathbb{R}^2 \to \mathbb{R}$ como

$$f(x, y) = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

donde $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = t(x, y)$. Entonces

$$\begin{aligned} f(x, y) &= \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_0^1 \left((tx + (ty)^2)x + 2(tx)(ty)y\right)dt \\ &= \int_0^1 \left(tx^2 + 3t^2y^2x\right)dt = x^2 \int_0^1 tdt + 3y^2x \int_0^1 t^2dt \\ &= \frac{1}{2}x^2 + y^2x \end{aligned}$$

■

Ejemplo 12. Consideremos el campo del ejemplo 4, $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + 4y, ax + y)$. En ese ejemplo se afirmó que $\mathbf{F}$ es conservativo si y sólo si $a = 4$. Ahora podemos quedar convencidos de este hecho apoyándonos en el teorema 7.4.4. En efecto, $\mathbf{F}$ es conservativo si y sólo si

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(x + 4y) = 4 = a = \frac{\partial}{\partial x}(ax + y) = \frac{\partial N}{\partial x}$$

Siendo entonces el campo $\mathbf{F}(x, y) = (x+4y, ax+y)$ conservativo, obtengamos una función potencial de él. Esta es $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda}$$

donde $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = t(x, y)$.Es decir

$$\begin{aligned} f(x, y) &= \int_0^1 \left((tx + 4ty)x + (4tx + ty)(y)\right)dt \\ &= \int_0^1 \left(tx^2 + 8txy + ty^2\right)dt = (x^2 + y^2 + 8xy) \int_0^1 tdt \\ &= \frac{1}{2}(x^2 + y^2 + 8xy) = \frac{1}{2}(x^2 + y^2) + 4xy \end{aligned}$$

como se había ya usado en el ejemplo 4. ■

Ejemplo 13. Consideremos el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, dado por $\mathbf{F}(x, y, z) = (3y^2z + ye^x, 6xyz + e^x, 3xy^2)$. Llamando M, N y P a las funciones coordenadas de $\mathbf{F}$, se tiene que

$$\begin{aligned} \frac{\partial M}{\partial y} &= \frac{\partial}{\partial y}(3y^2z + ye^x) = 6yz + e^x = \frac{\partial}{\partial x}(6xyz + e^x) = \frac{\partial N}{\partial x} \\ \frac{\partial M}{\partial z} &= \frac{\partial}{\partial z}(3y^2z + ye^x) = 3y^2 = \frac{\partial}{\partial x}(3xy^2) = \frac{\partial P}{\partial x} \\ \frac{\partial N}{\partial z} &= \frac{\partial}{\partial z}(6xyz + e^x) = 6xy = \frac{\partial}{\partial y}(3xy^2) = \frac{\partial P}{\partial y} \end{aligned}$$

Estando el campo definido en un conjunto convexo, concluimos que es conservativo. Hallemos una función potencial de $\mathbf{F}$. Definimos $f\colon \mathbb{R}^3 \to \mathbb{R}$ como la integral de $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}\colon [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = t(x, y, z)$. Entonces

$$
\begin{aligned}
f(x, y, z) &= \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} \\
&= \int_0^1 \big((3(ty)^2 tz + tye^{tx})x + (6(tx)(ty)(tz) + e^{tx})y + (3(tx)(ty)^2)z\big)dt \\
&= \int_0^1 \big(12t^3xy^2z + txye^{tx} + ye^{tx}\big)dt \\
&= 12xy^2z \int_0^1 t^3 dt + xy \int_0^1 te^{tx} dt + y \int_0^1 e^{tx} dt \\
&= 3xy^2z + xy\left(e^x\left(\frac{1}{x} - \frac{1}{x^2}\right) + \frac{1}{x^2}\right) + \frac{y}{x}(e^x - 1) = 3xy^2z + ye^x
\end{aligned}
$$

Es decir, la función $f(x, y, z) = 3xy^2z + ye^x$ es una función potencial del campo $\mathbf{F}$. Supongamos ahora que queremos calcular la integral del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\mu}\colon [0, \pi] \to \mathbb{R}^3$,

$$
\boldsymbol{\mu}(t) = \big(\operatorname{sen}(t/2) + \cos(2t) - 1,\ t\cos t,\ t^2 \operatorname{sen} t\big)
$$

el cual une el punto $\boldsymbol{\mu}(0) = (0, 0, 0)$ con el punto $\boldsymbol{\mu}(\pi) = (1, -\pi, 0)$. Ciertamente podríamos escribir la expresión que, según la definición dada en la sección anterior, nos conduce al valor de la integral deseada. Es decir, la expresión

$$
\int_0^{\pi} \big(F_1(x(t), y(t), z(t))x'(t) + F_2(x(t), y(t), z(t))y'(t) + F_3(x(t), y(t), z(t))z'(t)\big)\, dt
$$

donde $F_1(x, y, z) = 3y^2z + ye^x$, $F_2(x, y, z) = 6xyz + e^x$, $F_3(x, y, z) = 3xy^2$, $x(t) = \operatorname{sen}(t/2) + \cos(2t) - 1$, $y(t) = t\cos t$, $z(t) = t^2 \operatorname{sen} t$. Por supuesto que el simple imaginar la expresión que se obtendría al hacer las sustituciones correspondientes resulta una experiencia aterradora y más aterrador aún sería el intentar hacer las integrales coorespondientes. Sin embargo, podemos hacer una jugada más limpia (y mucho más fina) para calcular la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\mu}$. Sabemos que el campo $\mathbf{F}$ es conservativo. Por el inciso b) del teorema 7.4.1, el valor de la integral de $\mathbf{F}$ a lo largo del camino $\boldsymbol{\mu}$ sólo dependerá de los extremos de este camino y no del camino en sí. Es decir, si tomamos por ejemplo el camino $\boldsymbol{\nu}\colon [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\nu}(t) = (t, -\pi t, 0)$ el cual comparte los puntos inicial y final con $\boldsymbol{\mu}$ (pues $\boldsymbol{\mu}(0) = \boldsymbol{\nu}(0) = (0, 0, 0)$ y $\boldsymbol{\mu}(\pi) = \boldsymbol{\nu}(1) = (1, -\pi, 0)$), entonces la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\mu}$ será la misma que la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\nu}$. Este último camino es mucho más manejable que el camino dado $\boldsymbol{\mu}$. Se tendría entonces que

$$
\int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\nu}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_0^1 \big(-\pi te^t + e^t(-\pi)\big)\, dt = -\pi e
$$

Más aún, teniendo la función potencial de $\mathbf{F}$, $f(x, y, z) = 3xy^2z + ye^x$, podemos calcular la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\nu}$ como se hizo en el ejemplo 1 (usando el corolario del teorema 7.4.1), quedándonos

$$
\int_{\boldsymbol{\nu}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \big(3xy^2z + ye^x\big)_{\substack{x=1 \\ y=-\pi \\ z=0}} - \big(3xy^2z + ye^x\big)_{x=y=x=0} = -\pi e
$$

■

Si tenemos un campo conservativo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, de clase $\mathscr{C}^k$, $k \geq 1$, y queremos calcular la integral de línea de $\mathbf{F}$ a lo largo de un camino dado $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, según lo establecido en el teorema 7.4.1 el valor de esta integral depende solamente de los puntos extremos del camino $\boldsymbol{\lambda}$, $\mathbf{p} = \boldsymbol{\lambda}(a)$ y $\mathbf{q} = \boldsymbol{\lambda}(b)$. Así, tiene sentido una expresión del tipo

$$\int_{\mathbf{p}}^{\mathbf{q}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

para el cálculo de la integral de línea del campo conservativo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}$ que conecta el punto $\mathbf{p}$ (inicial) con el punto $\mathbf{q}$ (final). Al hacer este cálculo podremos tomar *cualquier* camino $\boldsymbol{\lambda}$ (con su imagen contenida en U) que empiece en $\mathbf{p}$ y termine en $\mathbf{q}$.

Ejemplo 14. Calculemos la integral

$$\int_{(1,0)}^{(3,2)} 2xy\,dx + x^2\,dy$$

Nótese que el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (2xy, x^2)$ es tal que

$$\frac{\partial}{\partial y}(2xy) = 2x = \frac{\partial}{\partial x}(x^2)$$

y, puesto que está definido en un conjunto convexo, podemos concluir que es un campo conservativo. Así tiene sentido calcular la integral de línea de $\mathbf{F}$ a lo largo de un camino (cualquiera) $\boldsymbol{\lambda}$ que conecte el punto $\mathbf{p} = (1, 0)$ con el punto $\mathbf{q} = (3, 2)$. Un camino semejante es, por ejemplo, $\boldsymbol{\lambda}: [1, 3] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t-1)$. Entonces

$$\begin{aligned} \int_{(1,0)}^{(3,2)} 2xydx + x^2dy &= \int_1^3 \left(2(t)(t-1)(t)' + (t)^2(t-1)'\right)dt \\ &= \int_1^3 (-2t + 3t^2)dt = \left[t^2 + t^3\right]_1^3 = -18 \end{aligned}$$ ■

Ejemplo 15. Calculemos la integral

$$\int_{(0,0,0)}^{(3,-2,5)} 3x\,dx + y^3\,dy - z^2\,dz$$

De nuevo vemos que se cumplen todas las condiciones del teorema 7.4.4 (todas las derivadas parciales son iguales a cero), por lo que el campo es conservativo. Tomemos entonces el camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (3t, -2t, 5t)$. La integral requerida es la integral de línea del campo $\mathbf{F}(x, y, z) = (3x, y^3, -z^2)$ a lo largo del camino $\boldsymbol{\lambda}$. Se tiene

$$\begin{aligned} \int_{(0,0,0)}^{(3,-2,5)} 3x\,dx + y^3\,dy - z^2\,dz &= \int_0^1 \left(3(3t)(3t)' + (-2t)^3(-2t)' - (5t)^2(5t)'\right)dt \\ &= \int_0^1 +(16t^3 - 125t^2 + 27t)\,dt \\ &= \left[4t^4 - \frac{125}{3}t^3 + \frac{27}{2}t^2\right]_0^1 = -\frac{115}{6} \end{aligned}$$ ■

Veamos ahora una alternativa al procedimiento planteado en los ejemplos 11, 12 y 13 anteriores para obtener funciones potenciales de campos conservativos. Haremos la discusión en el caso de campos en $\mathbb{R}^2$. Supongamos entonces que el campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (M, N)$ es conservativo y que queremos determinar una función potencial para él. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ la función potencial requerida. Se debe tener entonces que

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right) = \big(M(x, y), N(x, y)\big) = F(x, y)$$

Es decir, estamos procurando una función f que satisfaga el par de condiciones

$$\frac{\partial f}{\partial x}(x, y) = M(x, y), \qquad \frac{\partial f}{\partial y}(x, y) = N(x, y)$$

para (x, y) en U. Partiendo de la primera de estas expresiones, podemos decir que la función f que buscamos debe ser del tipo

$$f(x, y) = \int M(x, y)\, dx + g(y)$$

donde se realiza la integral respecto de la variable x, tomando a y como constante, y $g(y)$ es una función que depende solamente de y. Una función de este tipo satisface la primera de las condiciones requeridas para $f(x, y)$. Impongamos la segunda condición. Se tiene, derivando $f(x, y)$ respecto de y

$$\begin{aligned} N(x, y) &= \frac{\partial f}{\partial y}(x, y) = \frac{\partial}{\partial y}\Big(\int M(x, y) dx + g(y)\Big) \\ &= \frac{\partial}{\partial y}\int M(x, y) dx + g'(y) \end{aligned}$$

de donde se obtiene que la función $g(y)$ debe ser

$$g(y) = \int\left(N(x, y) - \frac{\partial}{\partial y}\int M(x, y) dx\right) dy$$

y entonces la función potencial requerida es

$$f(x, y) = \int M(x, y)\, dx + \int\left(N(x, y) - \frac{\partial}{\partial y}\int M(x, y) dx\right) dy$$

Este argumento que presenta la manera de construir la función potencial f del campo $\mathbf{F}$ queda concluido si mostramos que la función $g(y)$ ya establecida *es efectivamente una función que depende sólo de* y. Esto en principio no es evidente, pues en la fórmula que se obtuvo para determinar $g(y)$ "se ven muchas x metidas". Para ver que estas x son "ilusiones ópticas", derivemos respecto de x el integrando en la expresión para $g(y)$. Se tiene

$$\begin{aligned} \frac{\partial}{\partial x}\left(N(x, y) - \frac{\partial}{\partial y}\int M(x, y) dx\right) &= \frac{\partial N}{\partial x}(x, y) - \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\int M(x, y) dx\right) \\ &= \frac{\partial N}{\partial x}(x, y) - \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\int M(x, y) dx\right) \\ &= \frac{\partial N}{\partial x}(x, y) - \frac{\partial}{\partial y}\big(M(x, y)\big) = \frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y) = 0 \end{aligned}$$

donde la última igualdad se apoya en el hecho de que el campo **F** es conservativo, y por lo tanto $\frac{\partial N}{\partial x} = \frac{\partial M}{\partial y}$. Hemos probado entonces que el integrando en la expresión que define a $g(y)$ tiene derivada parcial respecto de x igual a cero, lo cual, junto con algunos detalles técnicos adicionales tomados de los teoremas del valor medio para derivadas, nos permite concluir que ese integrando *no contiene a la variable x*. Esto termina de probar entonces que tal expresión para $g(y)$, produce efectivamente una función que depende sólo de y. Veamos algunos ejemplos.

Ejemplo 16. Consideremos nuevamente el campo del ejemplo 11, $\mathbf{F}:\mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y^2, 2xy)$. Puesto que $\frac{\partial N}{\partial x} = \frac{\partial M}{\partial y}$, y $\mathbb{R}^3$ es convexo, el campo **F** es conservativo. Obtengamos una función potencial para **F**. Se quiere entonces una función $f:\mathbb{R}^2 \to \mathbb{R}$ tal que

$$\frac{\partial f}{\partial x}(x, y) = M(x, y) = x + y^2, \quad \frac{\partial f}{\partial y}(x, y) = N(x, y) = 2xy$$

De la primera de estas expresiones obtenemos que

$$f(x, y) = \int (x + y^2)\, dx + g(y) = \frac{1}{2}x^2 + y^2x + g(y)$$

Derivando respecto de y y usando la segunda de las condiciones establecidas para f se tiene

$$\frac{\partial f}{\partial y}(x, y) = 2xy + g'(y) = 2xy$$

de donde $g'(y) = 0$ y así, la función $g(y)$ es una constante, que podemos tomar como 0. Entonces una función potencial para el campo **F** es

$$f(x, y) = \frac{1}{2}x^2 + y^2x$$

como ya se había obtenido en el ejemplo 11. Pudimos haber partido de la segunda de las condiciones requeridas para f, es decir $\frac{\partial f}{\partial y}(x, y) = N(x, y) = 2xy$, quedándonos

$$f(x, y) = \int 2xy\, dy + h(x) = xy^2 + h(x)$$

Derivando respecto de x y usando la primera de las condiciones de f se obtiene

$$\frac{\partial f}{\partial x}(x, y) = y^2 + h'(x) = x + y^2$$

de donde $h'(x) = x$, y entonces $h(x) = \frac{1}{2}x^2$. Es decir, una función potencial de **F** es

$$f(x, y) = xy^2 + h(x) = xy^2 + \frac{1}{2}x^2$$

que es el mismo resultado ya obtenido. ∎

Ejemplo 17. Las ideas establecidas anteriormente para obtener funciones potenciales de campos conservativos en $\mathbb{R}^2$, se extienden fácilmente al caso $n = 3$. Veamos un ejemplo concreto. Retomemos el campo del ejemplo 13, $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, dado por $\mathbf{F}(x, y, z) = (3y^2z + ye^x, 6xyz + e^x, 3xy^2)$. Ya se vio que este campo es conservativo. Procuramos entonces una función $f: \mathbb{R}^3 \to \mathbb{R}$ de modo que

$$\operatorname{grad} f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = F(x, y, z)$$

es decir, que

$$\frac{\partial f}{\partial x} = 3y^2z + ye^x, \quad \frac{\partial f}{\partial y} = 6xyz + e^x, \quad \frac{\partial f}{\partial z} = 3xy^2$$

De la primera de estas expresiones se obtiene que

$$f(x, y, z) = \int (3y^2z + ye^x)\, dx + g(y, z) = 3y^2zx + ye^x + g(y, z)$$

Derivando esta expresión respecto de y y respecto de z y usando las dos expresiones restantes para f nos queda

$$\frac{\partial f}{\partial y} = 6yzx + e^x + \frac{\partial g}{\partial y} = 6xyz + e^x$$
$$\frac{\partial f}{\partial z} = 3y^2x + \frac{\partial g}{\partial z} = 3xy^2$$

de donde $\frac{\partial g}{\partial y} = \frac{\partial g}{\partial z} = 0$, y entonces $g(y, z)$ es una función constante, que podemos tomar como cero. Entonces una función potencial para $\mathbf{F}$ es

$$f(x, y, z) = 3y^2zx + ye^x$$

como ya habíamos establecido en el ejemplo 13. ■

Terminamos esta sección con algunos comentarios acerca de las condiciones más generales sobre las condiciones de igualdad entre las derivadas parciales de las funciones coordenadas del campo $\mathbf{F}$ establecidas en el teorema 7.4.2 que son, además de necesarias, suficientes para garantizar que el campo $\mathbf{F}$ sea conservativo. Como ya apuntábamos, la "forma" del dominio U (su estructura topológica) donde está definido $\mathbf{F}$ es fundamental para tener la afirmación recíproca del teorema 7.4.2. Ya se vio que si U es convexo, las condiciones mencionadas son suficientes. Sin embargo, esta suficiencia se tiene también con condiciones mucho más generales que la convexidad de U. La propiedad topológica que debe poseer U es la de "conexidad simple".

En el espacio $\mathbb{R}^2$, podemos pensar de manera intuitiva un conjunto simplemente conexo como un conjunto que "no tiene hoyos". De otro modo, si pensamos en el conjunto U que tenga por frontera una curva cerrada simple C (la imagen de un camino cerrado simple, tal como fue definido en la sección 1), este conjunto será simplemente conexo si "el interior" del conjunto que limita C pertenece a U.

Aunque de manera intuitiva podamos llegar a tener una idea más o menos clara de lo que significa la conexidad simple para un conjunto del plano, resulta sorprendente darse cuenta que tratar de precisar las ideas en él involucradas es una tarea nada trivial. De modo más preciso, llegar a establecer con exactitud lo que significa el "interior" de un conjunto limitado por una curva cerrada simple, es una labor que, en principio, tiene que ser apoyada por el célebre Teorema de la Curva de Jordan,

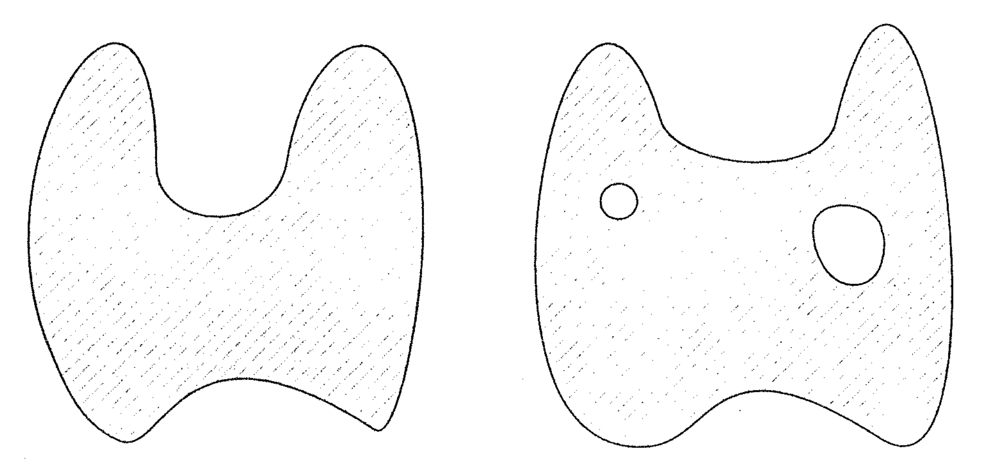


Figura 2. **a.** Conjunto simplemente conexo **b.** Conjunto no simplemente conexo.

el cual establece que "toda curva cerrada simple C en el plano, separa a éste en dos subconjuntos, uno acotado llamado *interior* a C, y otro no acotado llamado *exterior* a C, de los cuales la curva es la frontera común". Este hecho, que podría parecer "obvio" cuando se piensa en curvas cerradas simples como una circunferencia o una elipse (ver figura 3), es un resultado muy profundo de la parte de la matemática conocida como "Topología de Conjuntos". El nombre del teorema se debe al matemático francés Camile Jordan (1838-1922) quien fue de los primeros en precisar las ideas involucradas en el resultado, y dio una demostración (incompleta) de él. El Teorema fue probado por primera vez en 1905 por el matemático norteamericano Oswald Veblen (figura 3).

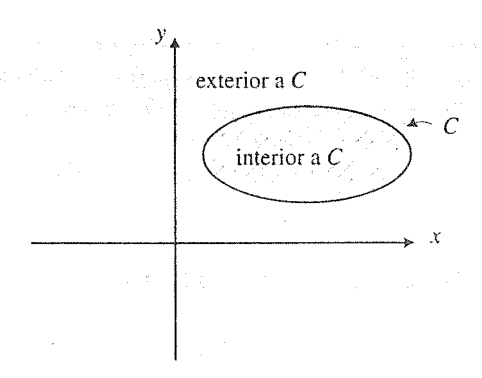


Figura 3. Teorema de la Curva de Jordan.

Nótese entonces que (en $\mathbb{R}^2$) un conjunto convexo es simplemente conexo, pero esta última propiedad es mucho más general que la de convexidad. La figura 4 ilustra estos hechos.

Enunciamos el resultado más general de esta sección, que mejora al teorema 7.4.4.

Teorema 7.4.5 Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ un campo de clase $\mathscr{C}^k$, $k \geq 1$, definido en el abierto simplemente conexo U de $\mathbb{R}^n$. Una condición necesaria y suficiente para

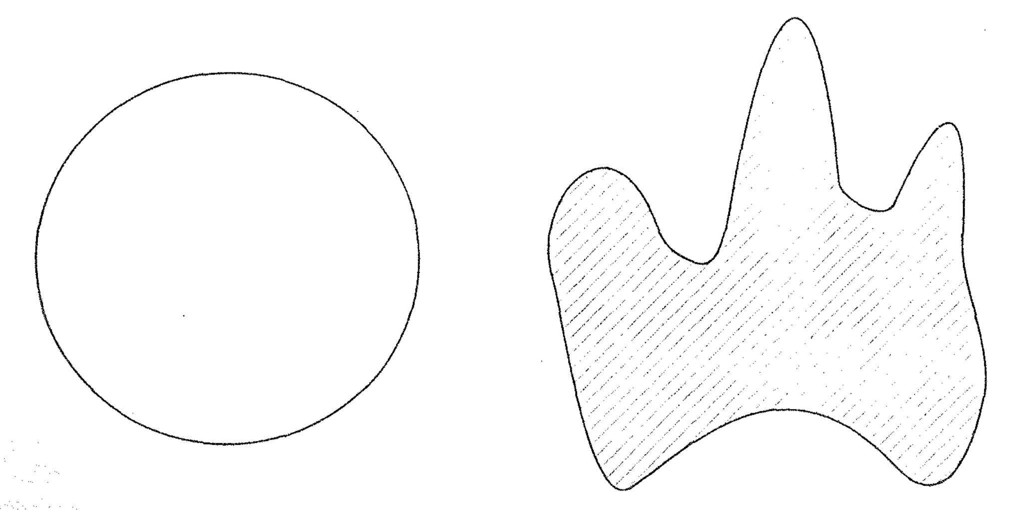


Figura 4. **a.** Conjunto convexo (y entonces, simplemente conexo) **b.** Conjunto simplemente conexo que no es convexo.

que el campo **F** sea conservativo es que

$$\frac{\partial F_i}{\partial x_j}(\mathrm{x}) = \frac{\partial F_j}{\partial x_i}(\mathrm{x})$$

para todo $x \in U$, $1 \leq i < j \leq n$. ■

La próxima sección es de estudio opcional y ha sido elaborada para los lectores que deseen tener una idea de los aspectos topológicos que presentan una exposición rigurosa de los temas aquí presentados. En un primer acercamiento a estos temas, el contenido de esta sección resulta más que suficiente para continuar estudiando la sección 8, donde se comienzan a abordar los aspectos físicos de la teoría de las integrales de línea desarrollada en este capítulo.

Ejercicios (Capítulo 7, Sección 4)

Para los campos $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ de los ejercicios 1–10:

a. Demuestre que es un campo conservativo.

b. Calcule la integral de línea de **F** a lo largo de *alguna* curva que una el origen de coordenadas con el punto (1, 1).

c. Determine una función potencial del campo **F**.

1. $\mathbf{F}(x, y) = (2xy + 1, x^2 + 4y)$

2. $\mathbf{F}(x, y) = (3y^3 + y + 2, 9xy^2 + x)$

3. $\mathbf{F}(x, y) = (2xy^2 + y + 5, 2x^2y + x + 2)$

4. $\mathbf{F}(x, y) = (2xy^3 + y + 1, 3x^2y^2 + x + 7)$

5. $\mathbf{F}(x, y) = (3x^2 + 2xy^2 + y, 2x^2y + x + 3y^2)$

6. $\mathbf{F}(x, y) = (15x^4 + 6x^2y^3 + y, 6x^3y^2 + x + 15y^4)$

7. $\mathbf{F}(x, y) = (y^2e^{x+y} + 1, ye^{x+y}(y + 2) + 1)$

8. $\mathbf{F}(x, y) = (4x + \operatorname{sen}^2 y, x \operatorname{sen} 2y + 1)$

9. $\mathbf{F}(x, y) = (y \cos x + \operatorname{sen} y + 1, \operatorname{sen} x + x \cos y + 1)$

10. $\mathbf{F}(x, y) = (e^x \operatorname{sen} y + 1, e^x \cos y + 1)$

Demuestre que cada uno de los campos $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ de los ejercicios 11–15, *no* es un campo conservativo: **a.** calculando la integral de línea de $\mathbf{F}$ a lo largo del círculo $x^2 + y^2 = 1$, **b.** usando del teorema 7.4.2.

11. $\mathbf{F}(x, y) = (xy, y^2)$

12. $\mathbf{F}(x, y) = (x + 4y, x - 5y)$

13. $\mathbf{F}(x, y) = (x^2 + y^2, xy)$

14. $\mathbf{F}(x, y) = (3xy, xy)$

15. $\mathbf{F}(x, y) = (x^2, xy^2)$

Para cada uno de los campos $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ de los ejercicios 16–20: **a.** Demuestre que es un campo conservativo. **b.** Calcule la integral de línea de $\mathbf{F}$ a lo largo de *alguna* curva que una el origen de coordenadas con el punto (1, 1, 1). **c.** Determine una función potencial del campo $\mathbf{F}$.

16. $\mathbf{F}(x, y, z) = (\pi, e, e^{\pi})$

17. $\mathbf{F}(x, y, z) = (2xy^3z, 3x^2y^2z, x^2y^3)$

18. $\mathbf{F}(x, y, z) = (e^{y+2z}, xe^{y+2z}, 2xe^{y+2z})$

19. $\mathbf{F}(x, y, z) = (3y + 2z^2, 3x, 4xz)$

20. $\mathbf{F}(x, y, z) = (z^2 + 1, 2z, 2xz + 2y)$

21. Calcular la integral de línea del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ dado por

$$F(x, y, z) = (e^{xz}(xyz^2 + yz), xze^{xz}, e^{xz}(x^2yz + xy))$$

a lo largo del camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^3$

$$\boldsymbol{\lambda}(t) = \left(\frac{\operatorname{senh} 5t^4}{\operatorname{senh} 5}, t^4 + 5t^3 - 3t^2 - 2t, \frac{1}{\ln 7} \ln(1 + 6t^8)\right)$$

y a lo largo del camino $\boldsymbol{\mu}: [0, 1] \to \mathbb{R}^3$

$$\boldsymbol{\mu}(t) = \left(\ln(t^2 - t + 1), \operatorname{sen}(t^3 + 3t^2 - 4t), \frac{\cosh(t^5 - t) - 1}{(t^2 + t + 1)^{4/7}}\right)$$

(Sugerencia: se supone que este ejercicio es fácil).

22. Demuestre que el campo $\mathbf{F}: \mathbb{R}^4 \to \mathbb{R}^4$, dado por

$$F(x, y, z, u) = (4uxy + 3yz, 2ux^2 + 3xz, 3xy + 3u^2z^2, 2x^2y + 2uz^3)$$

es conservativo. Calcule la integral de línea de este campo a lo largo de alguna curva que una el punto $(0, 1, 0, 1)$ con el punto $(1, 0, 1, 0)$. Determine una función potencial de este campo.

23. Demuestre que si los campos $\mathbf{F}, \mathbf{G}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, definidos en el conjunto abierto U de $\mathbb{R}^n$, son conservativos, entonces el campo $\mathbf{F} + \alpha\mathbf{G}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $(\mathbf{F} + \alpha\mathbf{G})(\mathbf{x}) = \mathbf{F}(\mathbf{x}) + \alpha\mathbf{G}(\mathbf{x})$, $\alpha \in \mathbb{R}$, es conservativo.

24. Sean $\varphi, \psi: I \subseteq \mathbb{R} \to \mathbb{R}$ dos funciones de clase $\mathscr{C}^1$ definidas en el intervalo abierto I de $\mathbb{R}$. Demuestre que

$$\oint_{\boldsymbol{\lambda}} \varphi(x)\, dx + \psi(y)\, dy = 0$$

donde $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ es cualquier camino cerrado cuya imagen está contenida en $|x|$.

25. Sea $f: \mathbb{R} \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ definida en $\mathbb{R}$. Demuestre que

$$\oint_{\boldsymbol{\lambda}} f(x^2 + y^2)(x dx + y\, dy) = 0$$

donde $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ es cualquier camino cerrado. Más en general, demuestre que la integral de línea del campo $\mathbf{F}: \mathbb{R}^n \to \mathbb{R}^n$ cuya i-ésima función coordenada es $F_i: \mathbb{R}^n \to \mathbb{R}$, $F_i(\mathbf{x}) = x_i f(\|\mathbf{x}\|)$, a lo largo de cualquier camino cerrado en $\mathbb{R}^n$, es igual a cero.

26. (Continuación del ejercicio 18 de la sección 3: un breve curso "hagalo usted mismo" de funciones de variable compleja. parte (iv) —y final—: el teorema de Cauchy).

y. Sea $f: U \subseteq \mathbb{C} \to \mathbb{C}$ una función holomorfa, definida en el conjunto abierto simplemente conexo U de $\mathbb{C}$. Sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ un camino cerrado de clase $\mathscr{C}^1$ cuya imagen está contenida en U, entonces (demuestre que)

$$\int_{\boldsymbol{\lambda}} f(z)\, dz = 0$$

Esta es una formulación del célebre teorema de Cauchy, uno de los resultados más importantes de la teoría de funciones de variable compleja. (Bosquejo de la demostración: el hecho de que f es holomorfa —de modo más preciso, el cumplimiento de las ecuaciones de Cauchy-Riemann— dice que los campos $\mathbf{F}$ y $\mathbf{G}$ según los cuales $\int_{\boldsymbol{\lambda}} f(z)\, dz = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} + i \int_{\boldsymbol{\lambda}} \mathbf{G} \cdot d\boldsymbol{\lambda}$ (ver definición de integral de f en el inciso s) del ejercicio 18 de la sección anterior), satisfacen las condiciones de igualdad de las derivadas parciales de las funciones coordenadas de estos campos establecidas en el teorema 7.4.5. Siendo U simplemente conexo, se concluye de este último teorema que los campos $\mathbf{F}$ y $\mathbf{G}$ son conservativos. La conclusión final se obtiene del teorema 7.4.1).

z. Revise los resultados de los incisos **t3**, **t4**, **w** y **x** del ejercicio 18 de la sección anterior, a la luz del teorema de Cauchy.

(*) 7.5 Un interludio topológico: conexidad

Esta es una sección cuyo estudio es de carácter opcional. Su contenido pertenece a la parte de la matemática conocida como Topología de Conjuntos, y el objetivo es tratar de *precisar* matemáticamente algunos de los detalles técnicos que surgieron en la sección anterior, todos ellos relacionados con una propiedad de los conjuntos de puntos en $\mathbb{R}^n$ llamada "conexidad". Nos tomamos la molestia de incursionar en estos temas por dos motivos fundamentales. El primero (y más importante) de ellos es para dejar la opción de que el estudiante interesado pueda avisorar a algunos de los temas que se estudian en esta parte de la matemática (temas que son elementales, desde la óptica de un curso de topología en una escuela de matemáticas; no tan elementales, a la óptica de un curso de cálculo como para el que está escrito este libro), y que se convenza de que estos temas son estudiados no solamente por el interés que ellos tienen por sí mismos, sino porque su comprensión nos ayuda a establecer con rigor las demostraciones de algunos resultados importantes que se estudian en el curso de cálculo (como los de la sección anterior). En segundo lugar, creemos que podría resultar muy formativo llegar a tener la sensación del (en ocasiones inmenso) trabajo que se tiene que desarrollar en matemáticas para hacer rigurosa una cierta idea, aunque ésta nos parezca en un primer acercamiento "obvia". Esto se puede lograr dando una mirada, aunque sea superficial, a las páginas siguientes.

En la sección 3 del capítulo 2, se introdujo el concepto de "conjunto abierto de $\mathbb{R}^n$". Recordemos: se dice que el conjunto $U \subseteq \mathbb{R}^n$ es abierto, si para cada $x_0 \in U$ existe una bola $B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x} - \mathbf{x}_0\| < r\}$ que está completamente contenida en U. Es decir, U es abierto si para cada $\mathbf{x}_0 \in U$ existe un $r > 0$ tal que $B(\mathbf{x}_0, r) \subset U$.

El que un conjunto U sea abierto o no, depende, entre otras cosas, del "ambiente" en el que está inmerso. Por ejemplo, tomemos U como la recta de los reales $\mathbb{R}$. Ciertamente U es abierto en $\mathbb{R}$, pues las bolas que se consideran en $\mathbb{R}$ son vecindades abiertas. Sin embargo, si vemos a U como el eje de las x en el plano xy (en $\mathbb{R}^2$), éste ya no será más un conjunto abierto. En efecto, dado un punto $P = (x_0, 0) \in U$, cualquier bola con centro en $\mathbf{p}$ y radio $r > 0$, contendrá puntos que no pertenecen a U.

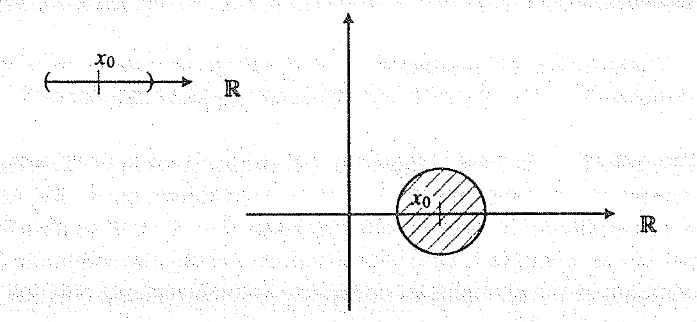


Figura 1. $U = \mathbb{R}$ es abierto en $\mathbb{R}$, pero U no es abierto en $\mathbb{R}^2$.

En nuestra definición de conjunto abierto, cuando decimos que "el conjunto $U \subseteq \mathbb{R}^n$ es abierto", sobreentendemos que queremos decir "el conjunto $U \subseteq \mathbb{R}^n$ es abierto en $\mathbb{R}^n$". En el estudio de la conexidad de conjuntos que haremos en esta sección, es necesario definir lo que significa que un conjunto U sea *abierto en otro conjunto X que lo contenga.* La idea para establecer esta definición es en esencia la misma que la de conjunto abierto en $\mathbb{R}^n$: para que $U \subseteq X$ sea abierto en X se requiere

que cada punto $\mathbf{x}_0 \in U$ sea centro de una bola $B(\mathbf{x}_0, r)$ tal que *la parte de esta bola que quede dentro de X, quede también dentro de U.* De modo más preciso, se tiene la siguiente definición.

Definición. Sea $U \subseteq X \subseteq \mathbb{R}^n$. Se dice que U es un conjunto abierto en X si para cada $\mathbf{x}_0 \in U$ existe $r > 0$ tal que $B(\mathbf{x}_0, r) \cap X \subset U$. ■

Nótese que si $X = \mathbb{R}^n$, se tiene $B(\mathbf{x}_0, r) \cap X = B(\mathbf{x}_0, r)$, y por lo tanto esta definición coincide con la dada previamente para conjuntos abiertos en $\mathbb{R}^n$.

Ejemplo 1. Todo conjunto $U \subseteq \mathbb{R}^n$ es abierto en él mismo. En efecto, dado $\mathbf{x}_0 \in U$, es claro que $B(\mathbf{x}_0, r) \cap U \subset U$ para cualquier $r > 0$. También es claro que el conjunto vacío $U = \varnothing$ es abierto en cualquier conjunto X de $\mathbb{R}^n$ (pues si no lo fuera, existiría un $\mathbf{x}_0 \in U$ tal que $B(\mathbf{x}_0, r) \cap X$ no está contenido en U para toda $r > 0$; pero tal $\mathbf{x}_0$ no existe, pues U es vacío). ■

Ejemplo 2. Sea X un conjunto abierto de $\mathbb{R}^n$ y sea $U \subseteq X$. El conjunto U es abierto en X si y sólo si U es abierto en $\mathbb{R}^n$. En efecto, supongamos que U es abierto en $\mathbb{R}^n$. Entonces dado $\mathbf{x}_0 \in U$ existe $r > 0$ tal que $B(\mathbf{x}_0, r) \subset U$. Como $B(\mathbf{x}_0, r) \cap X \subset B(\mathbf{x}_0, r)$, concluimos que dado $\mathbf{x}_0 \in U$ existe $r > 0$ tal que $B(\mathbf{x}_0, r) \cap X \subset U$, es decir, U es abierto en X. De forma recíproca, supongamos que U es abierto en X. Dado $\mathbf{x}_0 \in U$ existe $r_1 > 0$ de modo que $B(\mathbf{x}_0, r_1) \cap X \subset U$. Puesto que $\mathbf{x}_0 \in U \subseteq X$, y X es abierto en $\mathbb{R}^n$, existe un $r_2 > 0$ tal que $B(\mathbf{x}_0, r_2) \subset X$. Sea $r = \text{mín}(r_1, r_2)$, o bien, $B(\mathbf{x}_0, r) = B(\mathbf{x}_0, r_1) \cap B(\mathbf{x}_0, r_2)$. Entonces $B(\mathbf{x}_0, r) = B(\mathbf{x}_0, r_1) \cap B(\mathbf{x}_0, r_2) \subset B(\mathbf{x}_0, r_1) \cap X \subset U$. Es decir, dado $\mathbf{x}_0 \in U$, existe $r > 0$ de modo que $B(\mathbf{x}_0, r) \subset U$, lo que significa que U es abierto en $\mathbb{R}^n$. ■

Ejemplo 3. Sea $X = [-1, 1] \subset \mathbb{R}$. Cualquier subintervalo abierto $(a, b) \subset X$ es abierto en X (¿por qué?). El subconjunto $(0, 1] \subset X$ es abierto en X, aunque no abierto en $\mathbb{R}$. En efecto, si tomamos $x_0 \in (0, 1)$, no hay problema en verificar que haya una vecindad (una bola) con centro en x_0 y radio $r > 0$ que está completamente contenida en $(0, 1]$. Para $x_0 = 1$, tomemos $r < 1/2$, y observe que $B(1, 1/2) \cap [-1, 1] = (1/2, 1] \subset (0, 1]$. Así pues, $(0, 1]$ es abierto en $[-1, 1]$. ■

Diremos que el conjunto $U \subseteq X \subseteq \mathbb{R}^n$ es *cerrado en X*, si su complemento en X, es decir, el conjunto $X - U = \{x \in X, x \notin U\}$ es un conjunto abierto en X.

Ejemplo 4. En base al ejemplo 1, el conjunto vacío $\varnothing$ es cerrado en cualquier conjunto $X \subseteq \mathbb{R}^n$, puesto que su complemento $X - \varnothing = X$ es abierto en X. De modo análogo se ve que el conjunto X es cerrado en X, pues su complemento $X - X = \varnothing$ es abierto en X. Juntando estos resultados con los del ejemplo 1, concluimos que en un conjunto arbitrario $X \subseteq \mathbb{R}^n$, el conjunto X mismo y el conjunto $\varnothing$ son ejemplos de conjuntos simultáneamente abiertos y cerrados en X. ■

Ejemplo 5. Sean A y B dos conjuntos disjuntos abiertos en $\mathbb{R}^n$ y sea $X = A \cup B$. Obsérvese que X, siendo unión de dos conjuntos abiertos en $\mathbb{R}^n$, es un conjunto abierto en $\mathbb{R}^n$. Afirmamos que A es un subconjunto abierto y cerrado en X. En efecto, que A es abierto en X se deduce de lo mostrado en el ejemplo 2. En forma análoga se ve que B es abierto en X, y por lo tanto, siendo A el complemento en X del abierto B (es en este momento en que usamos que A y B son disjuntos), concluimos que A es cerrado en X. Un argumento similar nos muestra que B es también un subconjunto abierto y cerrado en X. ■

Ahora estamos en condiciones de presentar el concepto de conexidad.

7.5.1 Conjuntos conexos

De manera intuitiva, la propiedad de conexidad de un conjunto $X \subseteq \mathbb{R}^n$ nos dice si el conjunto X está constituído por una o varias piezas. Intuiremos que el conjunto X es conexo cuando X consta de una sola pieza, y no es conexo cuando está constituido por varias piezas. Por ejemplo, en el espacio $\mathbb{R}$, un intervalo es un conjunto conexo, pero el conjunto $X = [-1, 0] \cup [1, 3]$ no es conexo.

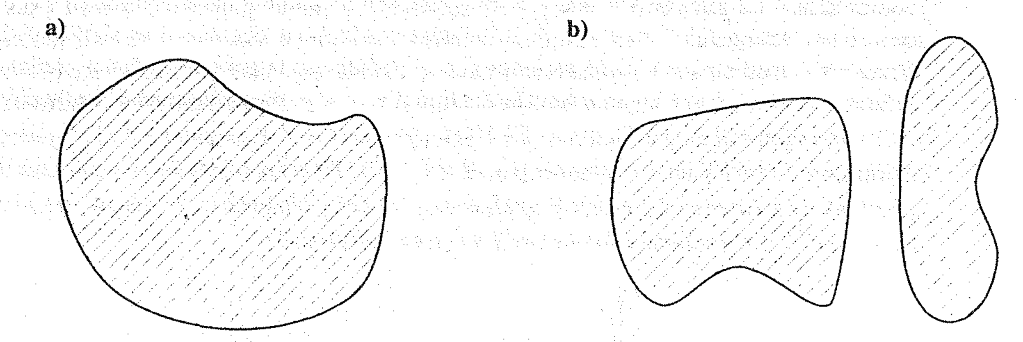


Figura 2. **a.** El conjunto X es conexo; **b.** El conjunto X no es conexo

En el ejemplo 1 se vio que en un conjuno arbitrario X, el mismo X y el conjunto vacío son abiertos en X. Además es claro que estos dos subconjuntos de X son disjuntos, pues $X \cap \varnothing = \varnothing$. Consideremos el problema de presentar al conjunto X como la unión de dos subconjuntos disjuntos (a una unión de este tipo se le llama *unión disjunta*) y abiertos en X. Este problema siempre tiene la solución (trivial) siguiente: $X = X \cup \varnothing$. Sin embargo, ésta no es, en general, la única solución. Consideremos, por ejemplo, el conjunto $X = (0, 1) \cup (2, 3)$ en $\mathbb{R}$. Es claro que X es un conjunto abierto de $\mathbb{R}$ y que, puesto que $(0, 1)$ y $(2, 3)$ también son abiertos en $\mathbb{R}$, lo son en X (ver ejemplo 2). Entonces, si queremos presentar a X como la unión de dos subconjuntos disjuntos abiertos en X, además de escribir $X = X \cup \varnothing$, podemos escribir también $X = A \cup B$, en donde $A = (0, 1)$, $B = (2, 3)$.

Definición. (Conjunto conexo). Se dice que el conjunto $X \subseteq \mathbb{R}^n$ es *conexo* si la única manera de escribir a X como la unión disjunta de dos subconjuntos abiertos en X, es la trivial, es decir, $X = X \cup \varnothing$. ■

Usando el ejemplo 4, el lector puede verificar que una definición equivalente de conexidad es la siguiente: el conjunto $X \subseteq \mathbb{R}^n$ es conexo si los únicos subconjuntos que son simultáneamente abiertos y cerrados en X son el mismo X y el vacío. El espacio $\mathbb{R}^n$ es un conjunto conexo. En $\mathbb{R}$, un conjunto $I \subseteq \mathbb{R}$ es conexo si y sólo si es un intervalo. La esfera en $\mathbb{R}^n$ de radio R

$$S_n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n | x_1^2 + x_2^2 + \cdots + x_n^2 = R^2\}$$

es un conjunto conexo. De todos estos hechos se pueden establecer demostraciones rigurosas, algunas de ellas basadas en resultados generales sobre el comportamiento de las imágenes de conjuntos conexos bajo funciones continuas. Quisiéramos, sin embargo, seguir usando nuestra idea intuitiva de conjunto conexo como conjunto formado por una sola pieza, AUNQUE... debemos ser cuidadosos con las sorpresas que suelen aparecer en matemáticas. Veamos el ejemplo siguiente.

Ejemplo 5. Consideremos el conjunto X de $\mathbb{R}^2$ dado por

$$X = \{(x, y)|y = \operatorname{sen}\frac{1}{x}, x > 0\} \cup \{(0, 0)\}$$

Así el conjunto X está formado por la unión disjunta de los conjuntos A y B, donde $A = \{(x, y)|y = \operatorname{sen}\frac{1}{x}, x > 0\}$, el cual es la mitad derecha de la gráfica de la función $y = \operatorname{sen}\frac{1}{x}$, y B es el origen de coordenadas. La gráfica de $y = \operatorname{sen}\frac{1}{x}$ es oscilante, con amplitud igual a la unidad y sus oscilaciones, que se van "replegando" en el eje de ordenadas conforme x decrece, van siendo cada vez "más frecuentes" a medida que x va decreciendo (i.e. a medida que la gráfica se acerca oscilando al eje y). De hecho, obsérvese que en un intervalo del tipo $0 < x < \epsilon$, para cualquier ϵ, la función $y = \operatorname{sen}\frac{1}{x}$ realiza *una infinidad de oscilaciones.* En efecto, para $0 < x < \epsilon$, se tiene $\frac{1}{x} > \frac{1}{\epsilon}$, y sabemos que sen θ realiza una infinidad de oscilaciones para $\theta \in (\frac{1}{\epsilon}, \infty)$. En otras palabras, en cualquier intervalo del tipo $(0, \epsilon)$, encontraremos parte de la gráfica de $y = sen\frac{1}{x}$ (de hecho, "la mayor" parte de la gráfica de $y = \operatorname{sen}\frac{1}{x}$). En esquema el conjunto X se ve en la (figura 3).

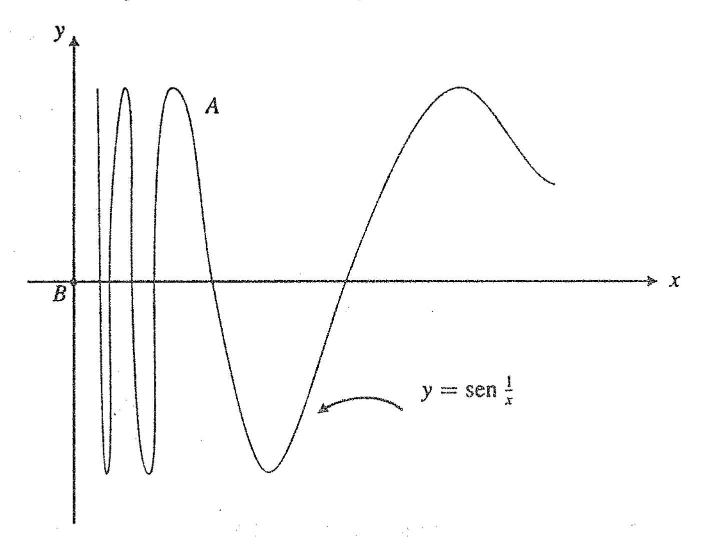


Figura 3. El conjunto X del ejemplo 5.

Nos disponemos a "probar" que el conjunto X *es conexo,* aunque esté formado por la unión disjunta de dos conjuntos A y B del plano $\mathbb{R}^2$. Si X no fuera conexo, debería haber dos conjuntos C y D disjuntos, abiertos en X, tales que $X = C \cup D$. La intuición nos dice que, como $X = A \cup B$, donde A y B son disjuntos, y cada uno de ellos consta de "una sola pieza"; la única posibilidad de que X no sea conexo es que los subconjuntos C y D fueran los subconjuntos A y B que definen a X. Lo que tendría que ocurrir entonces es que A y B debieran ser abiertos en X. Pero probaremos que esto no ocurre. Más precisamente, probaremos que el conjunto $B = \{(0, 0)\}$ NO ES ABIERTO EN X. Esto es consecuencia de que cualquier bola con centro en $(0, 0)$ y radio $r > 0$ *contiene varios pedazos de las infinitas oscilaciones que hace* $y = \operatorname{sen}\frac{1}{x}$ *cerca del origen*, de modo que no hay posibilidad alguna de que la intersección de X con una bola con centro en $(0, 0)$ y radio $r > 0$, esté completamente contenida en B (lo que equivale a afirmar que tal intersección sea solamente el origen de coordenadas). Esto sería posible sólo en caso de que el punto $(0, 0)$ estuviera "aislado" del

conjunto A, en el sentido de que podemos tener una bola alrededor de él que no comparte puntos con A. Esto muestra entonces que B no es abierto en X, de modo que aunque $X = A \cup B$, y $A \cap B = \varnothing$, ésta no es una manera de escribir a X como la unión disjunta de dos subconjuntos *abiertos en* X. Por supuesto que esto no muestra aún que X sea conexo, pero, como dijimos anteriormente, con un poco de intuición (y de buena voluntad), podemos aceptar que si ésta ($X = A \cup B$) no es la manera de escribir de modo no trivial a X como unión disjunta de subconjuntos abiertos en él, ya no hay más alternativas para hacerlo, quedándonos solamente la trivial $X = X \cup \varnothing$. ■

Estudiaremos ahora una propiedad relacionada con la conexidad, pero más restrictiva que ella.

7.5.2 Conjuntos conexos por caminos

La propiedad de los conjuntos que estudiaremos en este apartado se da en la definición siguiente

Definición. (Conjunto conexo por caminos). Se dice que el conjunto $X \subseteq \mathbb{R}^n$ es *conexo por caminos* si, dados dos puntos cualesquiera $\mathbf{p}, \mathbf{q} \in X$, hay un camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^n$, con $\boldsymbol{\lambda}([0, 1]) \subset X$, de modo que $\boldsymbol{\lambda}(0) = \mathbf{p}$, $\boldsymbol{\lambda}(1) = \mathbf{q}$. ■

Es decir, el conjunto $X \subseteq \mathbb{R}^n$ es conexo por caminos si dados dos puntos cualesquiera en X, éstos son los extremos de un camino $\boldsymbol{\lambda}$ cuya imagen está contenida en X. La pregunta natural ante esta nueva propiedad, es cuál es la relación con la conexidad del conjunto. El objetivo de esta sección es demostrar que una condición necesaria para que un conjunto sea conexo por caminos es que sea conexo, pero que tal condición no es suficiente. Para lograrlo, usaremos el resultado siguiente:

Lema ea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino en $\mathbb{R}^n$. La imagen $\boldsymbol{\lambda}([a, b]) \subseteq \mathbb{R}^n$ es un conjunto conexo.

Demostración. Llamemos X a la imagen $\boldsymbol{\lambda}([a, b]) \subseteq \mathbb{R}^n$. Nótese que ésta es la curva (traza) que describe el camino $\boldsymbol{\lambda}$ en $\mathbb{R}^n$. Para probar que X es conexo, supongamos que $X = A \cup B$, con A y B abiertos en X, $A \cap B = \varnothing$. Probaremos que uno de estos subconjuntos A o B es el vacío (y entonces, la única manera de escribir X como unión disjunta de subconjuntos abiertos en X es la trivial, es decir, X sería conexo). Sean $I = \boldsymbol{\lambda}^{-1}(A)$, $J = \boldsymbol{\lambda}^{-1}(B)$. Obsérvese que $[a, b] = I \cup J$ (en efecto, $t \in [a, b]$ si y sólo si $\boldsymbol{\lambda}(t) \in X = A \cup B$ si y sólo si $\boldsymbol{\lambda}(t) \in A$ o $\boldsymbol{\lambda}(t) \in B$ si y sólo si $t \in I$ o $t \in J$ si y sólo si $t \in I \cup J$). Además, $I \cap J = \varnothing$, pues en caso contrario existiría $t \in [a, b]$ de modo que $\boldsymbol{\lambda}(t) \in A \cap B$ (que se supone vacío). Probaremos que I y J son abiertos en $[a, b]$. Con esto terminaría la demostración, pues tendríamos $[a, b] = I \cup J$, con $I \cap J = \varnothing$, I y J abiertos en $[a, b]$, de donde, por ser $[a, b]$ conexo, uno de los dos subintervalos I o J es vacío, lo cual implicaría que A o B es vacío, lo cual, como vimos antes, implica la conexidad de X. Para terminar, veamos entonces que $I = \boldsymbol{\lambda}^{-1}(A)$ es abierto en $[a, b]$. Sea $t_0 \in I$ y consideremos el punto $\mathbf{s}_0 = \boldsymbol{\lambda}(t_0) \in A$. Puesto que A es abierto en X, existe una bola $B(\mathbf{s}_0, r) \subset \mathbb{R}^n$ tal que $B(\mathbf{s}_0, r) \cap X \subset A$. Por otro lado, como la función $\boldsymbol{\lambda}$ es continua (por definición) en t_0 (lo cual significa que $\lim_{t \to t_0} \boldsymbol{\lambda}(t) = \boldsymbol{\lambda}(t_0) = \mathbf{s}_0$. Ver la definición correspondiente en la sección 1 del capítulo 5), dado $r > 0$ existe $\delta > 0$ tal que

$$\text{si } t \in [a, b], \quad |t - t_0| < \delta \text{ entonces } \|\boldsymbol{\lambda}(t) - \mathbf{s}_0\| < r;$$

es decir, si $t \in (B(t_0, \delta) \cap [a, b])$ entonces $\boldsymbol{\lambda}(t) \in B(\mathbf{s}_0, r) \cap X \subset A$, lo cual se puede escribir como

$$\boldsymbol{\lambda}(B(t_0, \delta) \cap [a, b]) \subset B(\mathbf{s}_0, r) \cap X \subset A,$$

o bien como

$$B(t_0, \delta) \cap [a, b] \subset \boldsymbol{\lambda}^{-1}(B(s_0, r) \cap X) \subset \boldsymbol{\lambda}^{-1}(A) = I$$

En resumen, hemos demostrado que para un $t_0 \in I$ arbitrario, existe $\delta > 0$ de modo que $B(t_0, \delta) \cap [a, b] \subset I$. Esto prueba que I es abierto en $[a, b]$. La prueba de que J es abierto en $[a, b]$ es completamente análoga. ■

Probaremos entonces el resultado que nos dice que la conexidad por caminos implica la conexidad.

Teorema 7.5.1 Si el conjunto $X \subseteq \mathbb{R}^n$ es conexo por caminos, entonces es conexo.

Demostración. Fijemos un punto $\mathbf{p} \in X$. Para cada $\mathbf{q} \in X$, considere el camino $\boldsymbol{\lambda}_\mathbf{q}: [0, 1] \to \mathbb{R}^n$, de modo que $\boldsymbol{\lambda}_\mathbf{q}([0, 1]) \subset X$, $\boldsymbol{\lambda}_\mathbf{q}(0) = \mathbf{p}$, $\boldsymbol{\lambda}_\mathbf{q}(1) = \mathbf{q}$. Este camino existe, por las hipótesis hechas sobre el conjunto X. Sea $X_\mathbf{q} = \boldsymbol{\lambda}_\mathbf{q}([0, 1])$. Ciertamente se tiene que $X = \cup_{\mathbf{q}\in X} X_\mathbf{q}$. Para ver que X es conexo, supongamos que $X = A \cup B$, donde A y B son subconjuntos disjuntos abiertos en X. Como $\mathbf{p} \in X = A \cup B$, se tiene $\mathbf{p} \in A$ o $p \in B$. Supongamos que $\mathbf{p} \in A$. Se tiene, de hecho, $\mathbf{p} \in A \cap X_\mathbf{q}$ para toda $\mathbf{q} \in X$, pues es claro que, por la manera de definir el camino $\boldsymbol{\lambda}_\mathbf{q}$, $\mathbf{p} \in X_\mathbf{q}$ para toda $\mathbf{q}$. Puesto que $X_\mathbf{q} \subset X$, se tiene

$$X_\mathbf{q} = X \cap X_\mathbf{q} = (A \cup B) \cap X_\mathbf{q} = (A \cap X_\mathbf{q}) \cup (B \cap X_\mathbf{q})$$

Dejamos al lector que compruebe que los conjuntos $A \cap X_\mathbf{q}$ y $B \cap X_\mathbf{q}$ son subconjuntos disjuntos y abiertos en $X_\mathbf{q}$. Por el lema anterior, el conjunto $X_\mathbf{q} = \boldsymbol{\lambda}_\mathbf{q}([0, 1])$ = imagen del camino $\boldsymbol{\lambda}_\mathbf{q}$, es conexo. Entonces, de la expresión anterior para $X_\mathbf{q}$ se concluye que alguno de los dos subconjuntos $A \cap X_\mathbf{q}$ o $B \cap X_\mathbf{q}$ es vacío. Pero $\mathbf{p} \in A \cap X_\mathbf{q}$. Entonces se debe tener $B \cap X_\mathbf{q} = \varnothing$ para toda $\mathbf{q} \in X$. Entonces

$$B = B \cap X = B \cap \left(\cup_{\mathbf{q}\in X} X_\mathbf{q}\right) = \cup_{\mathbf{q}\in X}(B \cap X_\mathbf{q}) = \varnothing$$

Hemos entonces probado que siendo $X = A \cup B$, donde A y B son subconjuntos disjuntos abiertos en X, se tiene $B = \varnothing$. Esto muestra que X es conexo, como se quería. ■

La implicación recíproca del teorema anterior es falsa. El ejemplo más popular para ilustrar este hecho es con el conjunto

$$X = \{(x, y) | y = \operatorname{sen} \frac{1}{x}, x > 0\} \cup \{(0, 0)\}$$

estudiado en el ejemplo 5, donde se estableció que X es conexo. Sin embargo, es posible demostrar que este conjunto no es conexo por caminos. Aunque no entraremos en los detalles técnicos de la demostración de la afirmación anterior, es posible obtener cierto convencimiento de su validez si pensamos que para poder llegar de un punto $\mathbf{q}$ de la gráfica de la función $y = \operatorname{sen} \frac{1}{x}$, $x > 0$ al punto $\mathbf{p} = (0, 0) \in X$, con una función $f: [0, 1] \to \mathbb{R}^2$, cuya imagen esté en X, *ésta no puede ser continua*, pues como habíamos ya apuntado en el ejemplo 5, cerca del origen la función $y = \operatorname{sen} \frac{1}{x}$ oscila violentamente, y es claro que el origen queda siempre fuera de la gráfica de esta función. Este es pues un punto "inalcanzable" por un *camino* que conecte $\mathbf{p}$ con $\mathbf{q}$, cuya imagen esté siempre en X.

Para terminar este apartado veamos un resultado importante en el que se establecen las condiciones adicionales que se deben cumplir para que valga el recíproco del teorema 7.5.1.

Teorema 7.5.2 Si el conjunto abierto $X \subseteq \mathbb{R}^n$ es conexo, entonces es conexo por caminos.

Demostración. Sea $\mathbf{p} \in X$ un punto arbitrario pero fijo en X, y consideremos el conjunto

$$A = \{x \in X | \exists \boldsymbol{\lambda}: [0,1] \to \mathbb{R}^n,\ \boldsymbol{\lambda}([0,1]) \subset X,\ \boldsymbol{\lambda}(0) = \mathbf{p}, \boldsymbol{\lambda}(1) = \mathbf{x}\}$$

Este es entonces el conjunto de puntos $\mathbf{x} \in X$ que se pueden unir a $\mathbf{p}$ con un camino cuya imagen está contenida en X. Afirmamos que A es abierto en X. En efecto, tome $\mathbf{q} \in A$. Como $\mathbf{q} \in X$ y X es abierto en $\mathbb{R}^n$, existe una bola $B(\mathbf{q}, r)$ tal que $B(q, r) \subset X$. Se tiene de hecho $B(\mathbf{q}, r) \subset A$, pues si $\mathbf{w} \in B(\mathbf{q}, r)$, unimos a $\mathbf{w}$ con $\mathbf{q}$ con un camino rectilíneo (el cual queda contenido en la bola $B(\mathbf{q}, r)$ y por lo tanto contenido en X) y como $\mathbf{q}$ puede unirse con $\mathbf{p}$ con un camino contenido en X, hacemos la suma de estos dos caminos, la cual, a menos de una reparametrización, es un camino (contenido en X) definido en $[0, 1]$ que comienza en $\mathbf{p}$ y termina en $\mathbf{w}$. Esto significa entonces que $\mathbf{w} \in A$, y muestra que $B(\mathbf{q}, r) \subset A$. Así pues, el conjunto A es abierto en $\mathbb{R}^n$, y por lo tanto abierto en X (ver ejemplo 2). Por otra parte, el conjunto $B = X - A$, complemento del conjunto A, es decir, el conjunto de $x \in X$ para las cuales no existe un camino $\boldsymbol{\lambda}$ con su imagen contenida en X que conecte a $\mathbf{p}$ con $\mathbf{x}$, es también abierto en X. En efecto, dado $\mathbf{q} \in B$, por estar $\mathbf{q} \in X$ y X ser abierto, existe una bola $B(\mathbf{q}, r) \subset X$, la cual, de hecho, está contenida en B, pues si $\mathbf{w} \in B(\mathbf{q}, r)$, entonces w no se puede conectar a $\mathbf{p}$ con un camino contenido en X, ya que caso contrario, consideramos el camino que une a $\mathbf{p}$ con $\mathbf{w}$ y lo sumamos con el camino rectilíneo que une a $\mathbf{w}$ con $\mathbf{q}$. Así se tiene un camino contenido en X que conecta a $\mathbf{p}$ con $\mathbf{q}$, lo cual contradice que $\mathbf{q}$ pertenece a B. En resumen, $X = A \cup B$, donde A y B son subconjuntos disjuntos abiertos en X. Como X es conexo, uno de estos dos subconjuntos A o B es vacío. Puesto que el punto $\mathbf{p}$ pertenece a A, concluimos que B es vacío, y entonces $X = A$. Es decir, X es conexo por caminos. ■

Nótese que el teorema anterior se usó en la demostración de **2** $\Leftrightarrow$ **1** en el teorema 7.4.1 de la sección anterior. En él se definió una función potencial en el conjunto abierto U donde está definido el campo $\mathbf{F}$. Si este conjunto es conexo, entonces el teorema anterior asegura que es conexo por caminos, y en tal caso podemos tomar caminos contenidos en U que comiencen en un punto fijo $\mathbf{p}$ y terminen en un punto $\mathbf{x} \in U$, como se hizo en la demostración mencionada. Si el conjunto U no es conexo, entonces se puede demostrar (ya no entraremos en los detalles a este respecto) que U se puede presentar como unión disjunta de subconjuntos abiertos de él, los cuales son conexos. Estos subconjuntos de U se llaman "componentes conexas" y son "máximos" en el sentido siguiente: cada $x \in U$ pertenece a una componente conexa $C_{\mathbf{x}}$, la cual es el mayor subconjunto conexo de X que contiene a x (es claro que si X es conexo, $C_{\mathbf{x}} = X$ para toda $\mathbf{x} \in U$). Entonces, si en el teorema 7.4.1 el conjunto U no es conexo, definimos la función potencial f como se hizo en su demostración en cada una de las componentes conexas de U (que por ser conjuntos abiertos y conexos, son conexos por caminos).

7.5.3 Conjuntos simplemente conexos, homotopía

Los conjuntos que estudiaremos aquí son los del tipo más general, en donde es válido el recíproco del teorema 7.4.2 de la sección anterior. Es decir, en un conjunto de este tipo es donde debe estar definido un campo $\mathbf{F}$ para que la condición de igualdad entre las derivadas parciales de sus funciones componentes, sea una condición (además de necesaria) suficiente para que el campo sea, en ese conjunto donde está definido, conservativo.

La definición de la propiedad de conexidad simple de un conjunto se da en términos del concepto de homotopía, que estudiaremos ahora.

Definición. Sea $X \subseteq \mathbb{R}^n$ y sean $\boldsymbol{\lambda}, \boldsymbol{\mu}: [a, b] \to \mathbb{R}^n$ dos caminos cuya imagen está contenida en X (es decir, $\boldsymbol{\lambda}([a, b]) \subset X$, $\boldsymbol{\mu}([a, b]) \subset X$) que tienen los mismos puntos extremos (es decir, $\boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a)$, $\boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b)$). Se dice que estos caminos son *homotópicos* en X, si se da una función continua (llamada *homotopía entre* $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$), $H: [a, b] \times [0, 1] \to X$, de modo que

$$H(s, 0) = \boldsymbol{\lambda}(s), \quad H(s, 1) = \boldsymbol{\mu}(s), \quad \forall s \in [a, b] \tag{a}$$

$$H(a, t) = \boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a), \quad H(b, t) = \boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b), \quad \forall t \in [0, 1] \tag{b}$$

■

De manera intuitiva, se puede considerar una homotopía H como una deformación continua del camino $\boldsymbol{\lambda}$ al camino $\boldsymbol{\mu}$, la cual se efectúa dentro del conjunto X dado. En efecto, para cada $t_0 \in [0, 1]$ fija, podemos pensar en $H(s, t_0)$ como un camino en X, el cual, según las condiciones (b), tiene los mismos puntos extremos de $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$. Además, según las condiciones (a), si $t_0 = 0$, el camino $H(s, 0)$ es el camino $\boldsymbol{\lambda}$ y si $t_0 = 1$, el camino $H(s, 1)$ es el camino $\boldsymbol{\mu}$. De modo que mientras t recorre el intervalo $[0, 1]$, los caminos $H(s, t)$ pasan de manera continua del camino $\boldsymbol{\lambda}$ al camino $\boldsymbol{\mu}$, con su imagen siempre dentro del conjunto X, y manteniendo fijos los extremos. En esquema esto se ve así

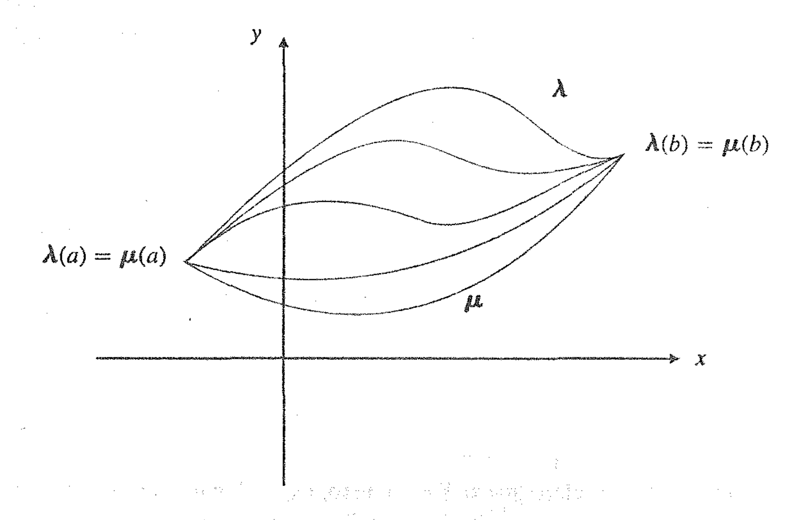


Figura 4. Homotopía entre el camino $\boldsymbol{\lambda}$ y el camino $\boldsymbol{\mu}$.

Un hecho importante de resaltar es que la relación de homotopía es una relación de equivalencia entre los caminos en X que conectan un punto $\mathbf{p} = \boldsymbol{\lambda}(a) \in X$ con un punto $q = \boldsymbol{\lambda}(b) \in X$. En efecto:

i. Es reflexiva: todo camino $\boldsymbol{\lambda}: [a, b] \to X$ es homotópico a sí mismo. La función $H: [a, b] \times [0, 1] \to X$, $H(s, t) = \boldsymbol{\lambda}(s)$ es una homotopía entre $\boldsymbol{\lambda}$ y $\boldsymbol{\lambda}$.

ii. Es simétrica: si el camino $\boldsymbol{\lambda}: [a, b] \to X$ es homotópico al camino $\boldsymbol{\mu}: [a, b] \to X$, digamos con la homotopía $H(s, t)$, entonces el camino $\boldsymbol{\mu}$ es homotópico al camino $\boldsymbol{\lambda}$, con la homotopía $\tilde{H}(s, t) = H(s, 1 - t)$.

iii. Es transitiva: si el camino $\boldsymbol{\lambda}: [a, b] \to X$ es homotópico al camino $\boldsymbol{\mu}: [a, b] \to X$, digamos con la homotopía $H_1(s, t)$, y el camino $\boldsymbol{\mu}: [a, b] \to X$ es homotópico al camino $\boldsymbol{\rho}: [a, b] \to X$,

digamos con la homotopía $H_2(s, t)$, entonces el camino $\boldsymbol{\lambda}$ es homotópico al camino $\boldsymbol{\rho}$, pues $\tilde{H}(s, t) = H_1(s, 2t)$, si $0 \leq t \leq 1/2$, $\tilde{H}(s, t) = H_2(s, 2t - 1)$, si $1/2 \leq t \leq 1$, es una homotopía entre $\boldsymbol{\lambda}$ y $\boldsymbol{\rho}$.

Entonces, el conjunto de caminos en X que conectan dos puntos dados de X, queda partido en clases de equivalencia con la relación de homotopía.

El resultado siguiente nos da condiciones suficientes para que dos caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ sean homotópicos en un conjunto dado $X \subseteq \mathbb{R}^n$.

Teorema 7.5.3 Sea $X \subseteq \mathbb{R}^n$ y sean $\boldsymbol{\lambda}, \boldsymbol{\mu}: [a, b] \to X$ dos caminos tales que $\boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a)$, $\boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b)$. Si para cada $s \in [a, b]$ el segmento de recta

$$[\boldsymbol{\lambda}(s), \boldsymbol{\mu}(s)] = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = t\boldsymbol{\mu}(s) + (1 - t)\boldsymbol{\lambda}(s), 0 \leq t \leq 1\}$$

queda contenido en X, entonces $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ son homotópicos.

Demostración. Defina $H: [a, b] \times [0, 1] \to \mathbb{R}^n$ como

$$H(s, t) = t\boldsymbol{\mu}(s) + (1 - t)\boldsymbol{\lambda}(s)$$

Esta es una función continua cuya imagen siempre está en X (por hipótesis). Además, $H(s, 0) = \boldsymbol{\lambda}(s)$, $H(s, 1) = \boldsymbol{\mu}(s)$, $\forall s \in [a, b]$, $H(a, t) = t\boldsymbol{\mu}(a) + (1 - t)\boldsymbol{\lambda}(a) = t\boldsymbol{\lambda}(a) + (1 - t)\boldsymbol{\lambda}(a) = \boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a)$, $H(b, t) = t\boldsymbol{\mu}(b) + (1 - t)\boldsymbol{\lambda}(b) = t\boldsymbol{\lambda}(b) + (1 - t)\boldsymbol{\lambda}(b) = \boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b)$, $\forall t \in [0, 1]$. Es decir, H es una homotopía entre $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$. ■

Las condiciones establecidas en el resultado anterior son suficientes, pero no necesarias, para que los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ sean homotópicos. Por ejemplo, considere el siguiente subconjunto de $\mathbb{R}^2$

$$X = \mathbb{R}^2 - \{(x, y) | x^2 + y^2 \leq 9\}$$

Es decir, X es el plano $\mathbb{R}^2$ al que se le ha eliminado la bola cerrada $x^2 + y^2 \leq 9$. En X definimos los dos caminos $\boldsymbol{\lambda}, \boldsymbol{\mu}: [0, 1] \to X$, dados por

$$\boldsymbol{\lambda}(s) = \left(5(2s^2 - 1), 7(1 - (2s^2 - 1)^2)\right)$$
$$\boldsymbol{\mu}(s) = \left(5(2s^{1/14} - 1), 5(1 - (2s^{1/14} - 1)^2)\right)$$

Para $s = 0.3$, se tiene $\boldsymbol{\lambda}(0.3) = (-4.1, 2.2932)$, $\boldsymbol{\mu}(0.3) = (4.17596, 1.51227)$. El segmento de recta que une estos dos puntos es

$$[\boldsymbol{\lambda}(0.3), \boldsymbol{\mu}(0.3)] = \{(8.27596t - 4.1, 2.2932 - 0.780929t), 0 \leq t \leq 1\}$$

el cual, pasa por el punto $(0, 1.90631)$ que claramente no pertenece a X. Sin embargo, los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ son homotópicos pues $H: [0, 1] \times [0, 1] \to X$,

$$H(s, t) = t\boldsymbol{\mu}(s^{14}) + (1 - t)\boldsymbol{\lambda}(\sqrt{s})$$

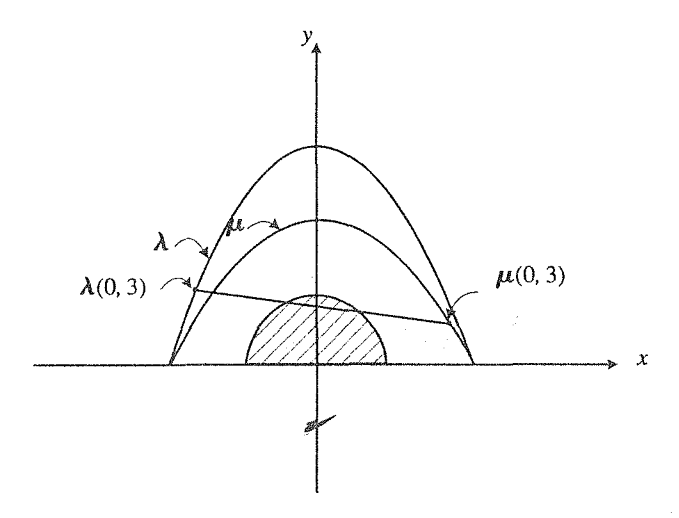


Figura 5. Los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ son homotópicos en el conjunto $X = \mathbb{R}^2 - \{x^2 + y^2 \leq 9\}$, a pesar de que el segmento de recta $[\boldsymbol{\lambda}(0,3), \boldsymbol{\mu}(0,3)]$ no está contenido en X.

es una homotopía entre ellos. Este ejemplo se ilustra gráficamente en la figura 5.

Por otra parte, nótese que si el conjunto $X \subseteq \mathbb{R}^n$ es convexo (es decir, si para cada $\mathbf{p}, \mathbf{q} \in X$, se tiene un segmento de recta que une $\mathbf{p}$ con $\mathbf{q} = [\mathbf{p}, \mathbf{q}] \subset X$), entonces las hipótesis del teorema anterior se cumplen automáticamente y por lo tanto cualquiera de los dos caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ definidos en $[a, b]$ con las mismas extremidades, serán homotópicos. De hecho, en este caso la función $H: [a, b] \times [0, 1] \to X$

$$H(s, t) = t\boldsymbol{\mu}(s) + (1 - t)\boldsymbol{\lambda}(s)$$

es una homotopía entre los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$, como el lector podrá verificar sin ningún problema. Así pues, en el espacio $\mathbb{R}^n$ o en una bola de este espacio, cualquiera de los dos caminos con las mismas extremidades son homotópicos.

Por ejemplo, en $\mathbb{R}^2$ consideremos los caminos $\boldsymbol{\lambda}, \boldsymbol{\mu}: [0, 1] \to \mathbb{R}^2$

$$\boldsymbol{\lambda}(s) = \big(5s - 2, -0.510571(5s - 2)^2 + 0.825217(5s - 2) + 5\big)$$
$$\boldsymbol{\mu}(s) = \big(5s - 2, (5s - 2)\cos(10s - 4)\big)$$

Estos caminos comparten sus puntos extremos $\boldsymbol{\lambda}(0) = \boldsymbol{\mu}(0) = (-2, 1.30728)$, $\boldsymbol{\lambda}(1) = \boldsymbol{\mu}(1) = (3, 2.88051)$, y, como $\mathbb{R}^2$ es convexo, son homotópicos. Una homotopía entre ellos es $H: [0, 1] \times [0, 1] \to \mathbb{R}^2$

$$\begin{aligned} H(s, t) &= t\boldsymbol{\mu}(s) + (1 - t)\boldsymbol{\lambda}(s) \\ &= \big(5s - 2, t(5s - 2)\cos(10s - 4) + (t - 1)(12.7642s^2 - 14.33755s - 1.30728)\big) \end{aligned}$$

La figura 6 muestra el proceso de deformación continua de $\boldsymbol{\lambda}$ a $\boldsymbol{\mu}$ dado por esta homotopía, tomada para t entre 0 y 1 en intervalos de 0.1.

Debemos insistir en que el hecho de que dos caminos $\boldsymbol{\lambda}, \boldsymbol{\mu}: [a, b] \to \mathbb{R}^n$ sean o no homotópicos, depende no solamente de los caminos mismos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$, sino también *del conjunto X en que se encuentren sus imágenes.* Entonces, dos caminos dados pueden ser o no ser homotópicos en un

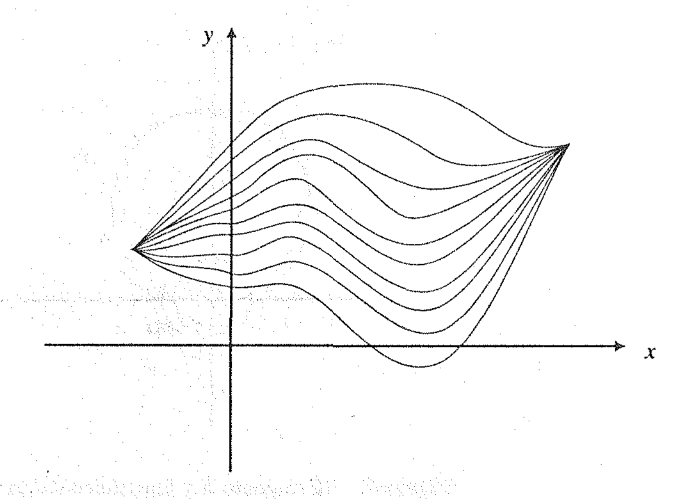


Figura 6. Proceso de deformación continua de $\boldsymbol{\lambda}$ a $\boldsymbol{\mu}$ dado por la homotopía H.

conjunto dado $X \subseteq \mathbb{R}^n$, y ellos mismos pueden ser o no ser homotópicos en otro conjunto y $\subseteq \mathbb{R}^n$. Por ejemplo, sea $X = \mathbb{R}^2 - \{(0,0)\}$. Consideremos los caminos $\boldsymbol{\lambda}, \boldsymbol{\mu}\colon [0,\pi] \to X$ dados por

$$\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t), \quad \boldsymbol{\mu}(t) = (\cos t, -\operatorname{sen} t)$$

El camino $\boldsymbol{\lambda}$ recorre la mitad superior del círculo unitario $x^2+y^2 = 1$, comenzando en $(1,0)$, mientras el camino $\boldsymbol{\mu}$ recorre la mitad inferior de este círculo. Estos caminos tienen el mismo punto inicial $(1,0)$ y el mismo punto final $(-1,0)$ *que no son homotópicos en* X. De manera intuitiva podemos pensar que no hay forma de pasar con un proceso continuo del semicírculo superior al semicírculo inferior con caminos cuya imagen está siempre en X, pues el origen de coordenadas no está incluido en este conjunto y, por lo tanto, los caminos intermedios tendrían que sufrir un rompimiento en algún momento (la demostración rigurosa de este hecho es muy técnica y preferimos omitirla). Sin embargo, en el mismo conjunto X consideremos el camino $\boldsymbol{\rho}$: $[0,\pi] \to X$,

$$\boldsymbol{\rho}(t) = \big(\cos(1.7048t) + 2\operatorname{sen}(1.7048t), \operatorname{sen}(1.7048t) - 2\cos(1.7048t) + 2\big)$$

Este es un camino que tiene los mismos extremos que $\boldsymbol{\lambda}$, pues $\boldsymbol{\rho}(0) = \boldsymbol{\lambda}(0) = (1,0)$ y $\boldsymbol{\rho}(\pi) = \boldsymbol{\lambda}(\pi) = (-1,0)$) y es fácil ver que con los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\rho}$ se cumplen las hipótesis del teorema 7.5.3, de modo que la función $H\colon [0,\pi] \times [0,1] \to X$, $H(s,t) = t\boldsymbol{\rho}(s) + (1-t)\boldsymbol{\lambda}(s)$ es una homotopía entre ellos. Así pues, en el conjunto $X = \mathbb{R}^2 - \{(0,0)\}$ los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\rho}$ son homotópicos, y los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ no lo son (ver la siguiente figura). Sin embargo, si tomamos ahora el conjunto $Y = \mathbb{R}^2 - \{(0,3)\}$, resultará que los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ anteriores serán homotópicos en Y, en tanto que los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\rho}$ no lo serán (¿por qué?) (ver figura 7).

El concepto de homotopía estudiado hasta este momento se aplica igualmente al caso en que los caminos $\boldsymbol{\lambda}$, $\boldsymbol{\mu}\colon [a,b] \to X \subseteq \mathbb{R}^n$ sean cerrados. En tal caso, puesto que en la definición de homotopía se requiere que los extremos de los caminos sean los mismos, se tendrá que $\boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a) = \boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b) = \mathbf{p}$. Diremos que el punto $\mathbf{p}$ es el "punto básico" de los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$. Así, diremos que los caminos cerrados $\boldsymbol{\lambda}$, $\boldsymbol{\mu}\colon [a,b] \to X \subseteq \mathbb{R}^n$ son homotópicos (en X) si se da una función continua $H\colon [a,b] \times [0,1] \to X$, de modo que $H(s,0) = \boldsymbol{\lambda}(s)$, $H(s,1) = \boldsymbol{\mu}(s)$, $H(a,t) = H(b,t) = \boldsymbol{\lambda}(a) = \boldsymbol{\lambda}(b) = \boldsymbol{\mu}(a) = \boldsymbol{\mu}(b)$, $\forall s \in [a,b]$, $\forall t \in [0,1]$. Cuando esta última

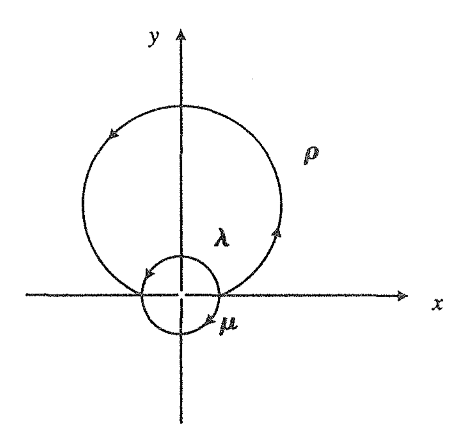


Figura 7. El conjunto X y los caminos λ, μ y ρ.

condición (que establece que el punto básico de los caminos cerrados λ y μ se mantiene fijo y para cada $t_0 \in I$ el camino $H(s, t_0)$ es cerrado) se modifica de manera que no se pida que el punto básico **p** se mantenga fijo, se llega a la noción de homotopía libre, cuya definición es la siguiente:

Definición. Sea $X \subseteq \mathbb{R}^n$ y sean $\lambda, \mu: [a, b] \to X$ dos caminos cerrados. Se dice que estos caminos son *libremente homotópicos* (en X) si hay una función continua $H: [a, b] \times [0, 1] \to X$ (llamada *homotopía libre*) de modo que

$$H(s, 0) = \lambda(s), \quad H(s, 1) = \mu(s), \quad \forall s \in [a, b] \tag{a}$$

$$H(a, t) = H(b, t), \quad \forall t \in [0, 1] \tag{b}$$

Así, una homotopía libre entre los caminos cerrados λ y μ, es una deformación continua de λ a μ (que, según las condiciones (a) comienza en λ y termina en μ) de manera que, según la condición (b), las "etapas intermedias" de la deformación siempre son caminos cerrados. En forma esquemática esto se ve en la figura siguiente.

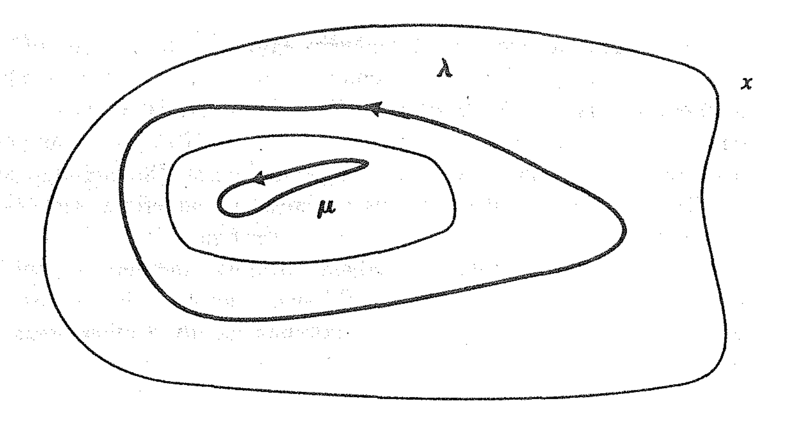


Figura 8. Los caminos λ y μ son libremente homotópicos en X.

Ejemplo 6. Considere los caminos $\boldsymbol{\lambda}, \boldsymbol{\mu}: [0, 2\pi] \to \mathbb{R}^2$,

$$\boldsymbol{\lambda}(s) = (r_1 \cos s, r_1 \operatorname{sen} s), \quad \boldsymbol{\mu}(s) = (r_2 \cos s, r_2 \operatorname{sen} s)$$

donde r_1 y r_2 son dos números positivos dados, digamos $r_1 > r_2$. Las imágenes de estos caminos son dos círculos concéntricos de radios r_1 y r_2. Es claro que $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ son libremente homotópicos. De hecho la función $H: [0, 2\pi] \times [0, 1] \to \mathbb{R}^2$,

$$\begin{aligned} H(s,t) &= t\boldsymbol{\mu}(s) + (1-t)\boldsymbol{\lambda}(s) \\ &= t(r_2 \cos s, r_2 \operatorname{sen} s) + (1-t)(r_1 \cos s, r_1 \operatorname{sen} s) \\ &= \big((tr_2 + (1-t)r_1)\cos s, (tr_2 + (1-t)r_1)\operatorname{sen} s\big) \end{aligned}$$

es una homotopía libre entre $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$. Nótese que los pasos intermedios de esta homotopía son círculos concéntricos de radio $(tr_2 + (1-t)r_1)$, $0 \leq t \leq 1$. Nótese también que este resultado sigue siendo válido si consideramos los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ con codominio en $\mathbb{R}^2 - \{(0,0)\}$ (figura 9).

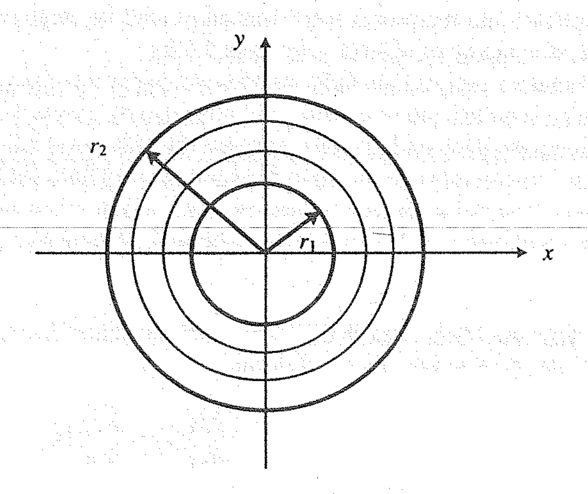


Figura 9. La homotopía libre entre los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ del ejemplo 6. ■

Con la noción de homotopía libre entre caminos cerrados es posible establecer el concepto de conjunto simplemente conexo.

Definición. (Conjunto simplemente conexo) Se dice que el conjunto $X \subseteq \mathbb{R}^n$ es *simplemente conexo*, si X es conexo por caminos y además todo camino cerrado $\boldsymbol{\lambda}: [a, b] \to X$ es libremente homotópico a un camino constante $\boldsymbol{\mu}: [a, b] \to X$, $\boldsymbol{\mu}(s) = \mathbf{p} \in X$. ■

El camino constante al que se refiere la definición anterior, es aquel cuya imagen consta solamente de un punto. Entonces un conjunto simplemente conexo es aquél en el cual todo camino cerrado se puede deformar de manera continua hasta "convertirlo" en un solo punto, de manera que todos los caminos en el proceso de deformación queden contenidos dentro del conjunto. En el caso de subconjuntos del espacio $\mathbb{R}^2$, la idea intuitiva que sugiere el concepto de conexidad simple es que

en un conjunto semejante no puede haber "hoyos", pues si los hubiera, cualquier camino cerrado que rodee a tales hoyos no se puede deformar continuamente *dentro del conjunto*, hasta convertirse en un punto. Esta idea intuitiva de ausencia de hoyos ya no funciona en el caso de subconjuntos de $\mathbb{R}^3$. Piense por ejemplo en una bola en $\mathbb{R}^3$ a la cual se ha eliminado una bola interna (como un durazno al que le hemos quitado el hueso); este es un conjunto con un hoyo, pero es simplemente conexo (¿por qué?). De cualquier modo advertimos que en general no es una tarea sencilla verificar que un conjunto es simplemente conexo. Quizás el ejemplo más claro de este tipo de conjuntos lo constituye el caso de los conjuntos *convexos*. En efecto, si $X \subseteq \mathbb{R}^n$ es convexo entonces este conjunto es simplemente conexo, pues cualquier camino cerrado $\boldsymbol{\lambda}: [a, b] \to X$ se deforma al punto $\mathbf{p}$ por medio de la homotopía libre $H: [a, b] \times [0, 1] \to X$, $H(s, t) = t\mathbf{p} + (1 - t)\boldsymbol{\lambda}(s)$. En particular entonces las bolas en $\mathbb{R}^n$ y todo el espacio $\mathbb{R}^n$ son ejemplos de conjuntos simplemente conexos.

Resumiendo lo discutido en esta sección, se tienen las siguientes implicaciones de las propiedades de un conjunto $X \subseteq \mathbb{R}^n$

$$\text{CONVEXO} \underset{(1)}{\Longrightarrow} \text{SIMPLEMENTE CONEXO} \underset{(2)}{\Longrightarrow} \text{CONEXO POR CAMINOS} \underset{(3)}{\Longrightarrow} \text{CONEXO}$$

Las implicaciones recíprocas son falsas en general (la implicación recíproca a (3) es verdadera si, además, el conjunto es abierto —teorema 7.5.2).

Terminamos esta sección indicando los resultados que conectan los conceptos de homotopía aquí estudiados (con los que se construye el concepto de conexidad simple), con la teoría de integrales de línea desarrollada en la sección anterior. El hecho más relevante de esta relación es el siguiente resultado, cuya demostración se puede encontrar en el libro [ELII]. (En el ejercicio 17 de la sección 9 se presenta una demostración de este teorema, según el teorema de Green que se estudiará en esa sección, en el caso $n = 2$, y cuando la homotopía H entre los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ es una función de clase $\mathscr{C}^2$.)

Teorema 7.5.4 Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un campo de clase $\mathscr{C}^k$, $(k \geq 1)$ definido en el abierto U de $\mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, que

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

para toda $x \in U$. Si $\boldsymbol{\lambda}, \boldsymbol{\mu}: [a, b] \to U$ son caminos seccionalmente $\mathscr{C}^1$, homotópicos en U (o bien, caminos cerrados seccionalmente $\mathscr{C}^1$, libremente homotópicos en U), entonces

$$\int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\mu}} F \cdot d\boldsymbol{\mu}$$ ■

Con este resultado tan importante, surge de inmediato el siguiente

Corolario ea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ un campo de clase $\mathscr{C}^k$, $(k \geq 1)$ definido en el conjunto abierto simplemente conexo U. Una condición necesaria y suficiente para que el campo $\mathbf{F}$ sea conservativo es que

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_j}{\partial x_i}(\mathbf{x})$$

para $x \in U$, $1 \leq i < j \leq n$.

Demostración. Que la condición es necesaria, se estableció en el teorema 7.4.2 en la sección anterior. Veamos que es suficiente. Sea $\boldsymbol{\lambda}: [a, b] \to U$ un camino cerrado en U, seccionalmente $\mathscr{C}^1$. Como U es simplemente conexo, $\boldsymbol{\lambda}$ es libremente homotópico a un camino constante $\boldsymbol{\mu}: [a, b] \to U$, $\boldsymbol{\mu}(s) = \mathbf{p}$. Por el teorema anterior, se tiene

$$\int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\mu}} F \cdot d\boldsymbol{\mu}$$

Pero es claro que la integral del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\mu}$ es cero ($\boldsymbol{\mu}'(s)$ es el vector cero). Por lo tanto

$$\int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} = 0$$

donde $\boldsymbol{\lambda}$ fue un camino cerrado en U arbitrario. Según el teorema 7.4.1, de la sección anterior, esta condición equivale a que el campo $\mathbf{F}$ sea conservativo. ■

Ejemplo 7. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right)$$

En el ejemplo 11 de la sección 3 se encontró que

$$\int_{\boldsymbol{\lambda}_k} \mathbf{F} \cdot d\boldsymbol{\lambda}_k = k\pi$$

donde $\boldsymbol{\lambda}_k$ es el camino $\boldsymbol{\lambda}_k: [0, k\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_k(t) = (r\cos t, r\operatorname{sen} t)$, $(0 < k \leq 2)$. En particular se tiene, con $k = 2$

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = 2\pi$$

Ya se había observado en la sección anterior que este campo $\mathbf{F} = (M, N)$ cumple con las condiciones de igualdad de las derivadas parciales de sus funciones coordenadas

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}\left(\frac{-y}{x^2 + y^2}\right) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) = \frac{\partial N}{\partial x}$$

Entonces, según el teorema 7.5.4, la integral de este campo a lo largo de cualquier curva libremente homotópica a $\boldsymbol{\lambda}$ debe ser 2π. De hecho, nótese cómo la traza de $\boldsymbol{\lambda}$ es un círculo *de radio* r y el resultado de la integral *no depende de* r. Como verificamos en el ejemplo 6, todos los círculos con centro en el origen son libremente homotópicos. Así, a la luz del teorema 7.5.4, se entiende claramente por qué tal valor de la integral no depende del radio del círculo concreto sobre el que se efectúa la integral. ■

Ejercicios (Capítulo 7, Sección 5)

Diga si los subconjuntos U de $\mathbb{R}^2$ dados en los ejercicios 1–10, son: **a** convexos, **b** conexos, **c** conexos por caminos, **d** simplemente conexos.

1. $U = \{(x, y)|x \leq 0\}$
2. $U = \{(x, y)|xy > 0\}$
3. $U = \{(x, y)|xy \geq 0\}$
4. $U = \{(x, y)|xy \geq 1\}$
5. $U = \{(x, y)|xy \leq 1\}$
6. $U = \{(x, y)|x^2 + y^2 \leq 1\}$
7. $U = \{(x, y)|x^2 + y^2 \geq 1\}$
8. $U = \{(x, y)|x^2 + y^2 \leq b, x^2 + y^2 \geq a\}(a < b)$
9. $U = \{(x, y)|x^2 + y^2 \leq b, x^2 + y^2 \geq a\}(a > b)$
10. $U = \{(x, y)|y = \operatorname{sgn} x\}$
11. Sean K_1 y K_2 dos conjuntos convexos en $\mathbb{R}^n$. ¿Es la unión $K = K_1 \cup K_2$ un conjunto convexo? Si lo es, demuéstrelo; en caso contrario, dé un contraejemplo.
12. Sean K_1 y K_2 dos conjuntos convexos en $\mathbb{R}^n$. Demuestre que la intersección $K = K_1 \cap K_2$ es un conjunto convexo. Aún más, pruebe que la intersección de una familia arbitraria de conjuntos convexos es un conjunto convexo.
13. (Envolventes convexas). Sea U un conjunto arbitrario de $\mathbb{R}^n$. La *envolvente convexa de* U, denotada por $C(U)$ se define como la intersección de todos los conjuntos convexos $K_i \subseteq \mathbb{R}^n$ que contienen a U. De hecho $C(U)$ es un conjunto convexo en $\mathbb{R}^n$ (¿por qué?).
    - **a.** Demuestre que si U es convexo, entonces $C(U) = U$.
    - **b.** Sea U el conjunto de $\mathbb{R}$ formado por los puntos $\mathbf{p}$ y $\mathbf{q}$ (digamos que $\mathbf{p} < \mathbf{q}$). Es decir $U = \{\mathbf{p}, \mathbf{q}\}$. Demuestre que la envolvente convexa de U es el intervalo cerrado $[\mathbf{p}, \mathbf{q}]$. ¿Cuál es la envolvente convexa del conjunto $U = \{\mathbf{p}, \mathbf{q}, 0.5(\mathbf{p} + \mathbf{q})\} \subset \mathbb{R}$?
    - **c.** Sea U la región en $\mathbb{R}^2$ limitada por un triángulo cuyos vértices son los puntos A, B y C. Verifique que la envolvente convexa de U es el círculo circunscrito al triángulo.
    - **d.** Demuestre que la envolvente convexa de un conjunto $U \subseteq \mathbb{R}^n$ está formada por todas las combinaciones lineales $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k$, en las que $\mathbf{x}_i \in U$, $\alpha_i \geq 0$, $i = 1, 2, \ldots, k$, $\alpha_1 + \alpha_2 + \cdots + \alpha_k = 1$.
14. En este ejercicio se probará que la imagen de un conjunto conexo bajo una función continua es un conjunto conexo. De modo más preciso, sea $\mathbf{f}: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función continua, y sea $V \subseteq U$ un subconjunto conexo de U. Entonces (probaremos que) $f(V) \subseteq \mathbb{R}^m$ es conexo. Suponga, para obtener una contradicción, que $f(V)$ no es conexo. Entonces ocurren A y B, subconjuntos abiertos disjuntos de $f(V)$ por los que $f(V) = A \cup B$.
    - **a.** Pruebe que los conjuntos $f^{-1}(A) = \{x \in V|f(x) \in A\}$, $f^{-1}(B) = \{x \in V|f(x) \in B\}$ son abiertos en V.
    - **b.** Pruebe que $V = f^{-1}(A) \cup f^{-1}(B)$, y que $f^{-1}(A) \cap f^{-1}(B) = \varnothing$.
    - **c.** Concluya que V no es conexo, lo cual contradice la hipótesis hecha sobre V.

15. En este ejercicio probaremos que si $I \subseteq \mathbb{R}$ es un subconjunto conexo de $\mathbb{R}$, entonces I debe ser un intervalo. Suponga, para obtener una contradicción, que la afirmación dada es falsa. Entonces debe haber puntos $a, b \in I$, $x_0 \notin I$, tales que $a < x_0 < b$. Sea $A = \{x \in \mathbb{R} | x < x_0\}$, $B = \{x \in \mathbb{R} | x > x_0\}$. Demuestre que $A \cap I$ y $B \cap I$ son conjuntos no vacíos, abiertos en I y disjuntos, cuya unión es I. Concluya entonces que I no es conexo. (La afirmación recíproca: "un intervalo I de $\mathbb{R}$ es conexo", también es verdadera, pero su demostración contiene algunos detalles técnicos sobre supremos e ínfimos de conjuntos de números reales que no presentamos. El lector interesado puede consultar, por ejemplo, la referencia [ELI]).

16. Tome el resultado del ejercicio anterior para demostrar el célebre Teorema del Valor Intermedio: sea $f: [a, b] \to \mathbb{R}$ una función continua definida en el intervalo compacto $[a, b]$. Entonces f toma *todos* los valores entre $f(a)$ y $f(b)$. En forma más precisa, si $y_0 \in \mathbb{R}$ es un número entre $f(a)$ y $f(b)$, existe (al menos un) $x_0 \in [a, b]$ tal que $f(x_0) = y_0$.

17. Con base en el resultado del ejercicio 14 haga una demostración (¡de una línea!) del lema que probamos en esta sección: sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino en $\mathbb{R}^n$. La imagen $\boldsymbol{\lambda}([a, b]) \subseteq \mathbb{R}^n$ es un conjunto conexo.

18. Sea $X = \{(x, y) | x^2 + y^2 \geq 1\} \subseteq \mathbb{R}^2$. Constate que X no es un conjunto simplemente conexo. Dé ejemplos de caminos en X que sean homotópicos y de caminos en X que no sean homotópicos.

19. Dé ejemplos de subconjuntos X del espacio $\mathbb{R}^3$ que no sean simplemente conexos.

20. En el ejemplo 10 de la sección 3 se calculó la integral de línea del campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + 4y, ax + y)$, a lo largo del círculo $x^2 + y^2 = r^2$, recorrido una vez en sentido antihorario. El resultado fue $\pi r^2(a - 4)$. Obsérvese que este resultado sí depende del círculo concreto sobre el que se esté integrando. Explique por qué este hecho no contradice el resultado del teorema 7.5.4, a pesar de que sabemos que todos los círculos con centro en el origen son libremente homotópicos entre sí. ¿Qué sucede si $a = 4$?

(**) **21.** Sean U y V dos conjuntos en $\mathbb{R}^n$ simplemente conexos. Suponga que su intersección $U \cap V$ es un conjunto conexo. Demuestre que su unión $U \cup V$ es un conjunto simplemente conexo.

(*) 7.6 Ecuaciones diferenciales exactas

La teoría desarrollada en la sección 4 de este capítulo tiene una traducción casi inmediata en una parte del estudio de las ecuaciones diferenciales. En esta sección, que es opcional (pues cae fuera del objeto de estudio de este libro), abordaremos el estudio de algunas ecuaciones diferenciales de primer orden, llamadas exactas, y veremos que la herramienta desarrollada en la sección 4 nos da las ideas fundamentales para resolver los problemas que se encuentran en esta parte de la matemática.

Recordemos rápidamente algunas definiciones básicas importantes: una *ecuación diferencial (ordinaria)* —que abreviaremos ED— es una ecuación que involucra una función incógnita determinada $y = g(x)$ (con propiedades de diferenciabilidad), su variable independiente x, y su(s) derivada(s). El *orden* de la ecuación es el de la derivada de mayor orden que en ella aparece. Entonces, una ED de primer orden es una expresión del tipo $F(x, y, y') = 0$, donde $y = g(x)$ es la función incógnita de la ecuación. Una *solución* de la ED de primer orden $F(x, y, y') = 0$, es una función derivable $y = y(x)$ que satisface la ecuación, es decir, tal que al sustituir esta función en la ecuación, reduce a ésta última a una identidad. O sea, $y = y(x)$ es solución de $F(x, y, y') = 0$, si $F(x, y(x), y'(x)) \equiv 0$

(para toda x en el dominio de $y(x)$, de modo que $(x, y(x), y'(x))$ esté en el dominio de F). La ED de primer orden $F(x, y, y') = 0$ se dice ser *resuelta respecto de la derivada*, si podemos despejar de ella a y' y escribir la ecuación como $y' = f(x, y)$. A continuación, consideraremos solamente ecuaciones diferenciales de primer orden resueltas respecto de la derivada, las cuales seguimos abreviando simplemente como ED.

Por el resto de esta sección, escribiremos la ED $y' = f(x, y)$ como

$$M(x, y)dx + N(x, y)dy = 0 \tag{E}$$

(donde se entiende que dy, la diferencial de la función incógnita $y = y(x)$, es $dy = y'\,dx$, y $f(x, y) = -\frac{M(x,y)}{N(x,y)}$). Supondremos que $M = M(x, y)$ y $N = N(x, y)$ son funciones de clase $\mathscr{C}^k$, $k \geq 0$, definidas en algún subconjunto abierto U de $\mathbb{R}^2$.

La ED (E) se dice ser *exacta* si existe una función $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$, de clase $\mathscr{C}^{k+1}$, de modo que la diferencial total de esta función (es decir, $df(x, y) = \frac{\partial f}{\partial x}(x, y)\,dx + \frac{\partial f}{\partial y}(x, y)\,dy$) sea

$$df(x, y) = M(x, y)\,dx + N(x, y)\,dy$$

Obsérvese que en el caso de que (E) sea una ED exacta, la ecuación se puede escribir como

$$df(x, y) = 0$$

de modo que la familia monoparamétrica de curvas en el plano $f(x, y) = c$ (donde c es una constante arbitraria) resulta ser solución de la ecuación (llamada *solución general*). En efecto, las funciones $y = y(x)$ definidas implícitamente por la expresión $f(x, y) = c$ (en los puntos $(x, y) \in U$ donde $\frac{\partial f}{\partial y}(x, y) \neq 0$, ver la sección 4 del capítulo 3, teorema 3.4.1) tienen por derivada a

$$y'(x) = -\frac{\dfrac{\partial f}{\partial x}(x, y)}{\dfrac{\partial f}{\partial y}(x, y)}$$

o bien, en términos de diferenciales

$$dy = -\frac{\dfrac{\partial f}{\partial x}(x, y)}{\dfrac{\partial f}{\partial y}(x, y)}\,dx$$

de donde

$$df(x, y) = \frac{\partial f}{\partial x}(x, y)\,dx + \frac{\partial f}{\partial y}(x, y)\,dy = 0$$

Ejemplo 1. La ecuación $2xy\,dx + (x^2 + 1)\,dy = 0$ es exacta, pues la función $f(x, y) = (x^2 + 1)y$ es tal que $df(x, y) = 2xy\,dx + (x^2 + 1)\,dy$. Entonces la solución general de la ecuación es $(x^2 + 1)y = c$. ■

Las preguntas obvias que surgen de la discusión anterior son: ¿cómo saber cuando una ED es exacta? En caso de que lo sea ¿cómo determinar su solución general? Estas preguntas tienen su respuesta en la sección 4 de este capítulo, si atendemos al paralelismo entre el estudio de las

ecuaciones diferenciales exactas y los campos conservativos. Este paralelismo lo encontramos asociando a la ED (E) el campo vectorial $F: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (M(x, y), N(x, y))$. Nótese entonces que la propiedad de exactitud de la ecuación (E) es equivalente a la propiedad del campo **F** de ser conservativo. En efecto, la ecuación (E) es exacta si y sólo si existe una función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ de clase $\mathscr{C}^{k+1}$ de modo que $df(x, y) = M(x, y)\,dx + N(x, y)\,dy$, es decir,

$$\frac{\partial f}{\partial x}(x, y) = M(x, y), \quad \frac{\partial f}{\partial y}(x, y) = N(x, y)$$

Esta función f resulta entonces una función potencial para el campo **F**, lo cual equivale a afirmar que este campo es conservativo. Así, podemos establecer el siguiente resultado, el cual, aunque está redactado en términos de ecuaciones diferenciales exactas, es copia textual de los teoremas ya estudiados en la sección 4 de este capítulo.

Teorema 7.6.1 Considere la ecuación diferencial

$$M(x, y)\,dx + N(x, y)\,dy = 0$$

donde $M, N: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son funciones de clase $\mathscr{C}^k$, $k \geq 0$, definidas en el conjunto abierto U de $\mathbb{R}^2$.

(I) Las siguientes afirmaciones son equivalentes:

a. La ecuación es exacta.

b. La integral de línea

$$\int_{\boldsymbol{\lambda}} M(x, y)\,dx + N(x, y)\,dy$$

sobre el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, que $\boldsymbol{\lambda}([a, b]) \subset U$, seccionalmente $\mathscr{C}^1$, depende solamente de los puntos extremos $\boldsymbol{\lambda}(a)$ y $\boldsymbol{\lambda}(b)$.

c. La integral de línea

$$\oint_{\boldsymbol{\lambda}} M(x, y)\,dx + N(x, y)\,dy$$

sobre cualquier camino cerrado $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, que $\boldsymbol{\lambda}([a, b]) \subset U$, seccionalmente $\mathscr{C}^1$, es igual a cero.

En estas circunstancias, la solución general de la ecuación se puede determinar como $f(x, y) = c$, donde

$$f(x, y) = \int_{\mathbf{p}}^{\mathbf{q}} M(x, y)\,dx + N(x, y)\,dy$$

con $\mathbf{p} = (x_0, y_0) \in U$, un punto arbitrario pero fijo, y $\mathbf{q} = (x, y)$.

(II) Si $k \geq 1$, una condición necesaria para que la ecuación diferencial sea exacta es que

$$\frac{\partial M}{\partial y}(x, y) = \frac{\partial N}{\partial x}(x, y), \quad \forall (x, y) \in U$$

(III) Si $k \geq 1$ y el conjunto abierto U donde están definidas las funciones M y N es simplemente conexo, la condición del apartado anterior también es suficiente para que la ecuación sea exacta. En este caso, la sulución general de la ecuación se puede escribir como

$$\int M(x, y)\,dx + \int \left(N(x, y) - \frac{\partial}{\partial y} \int M(x, y)\,dx\right) dy = c$$

o bien como

$$\int N(x, y)\,dy + \int \left(M(x, y) - \frac{\partial}{\partial x} \int N(x, y)\,dy\right) dx = c$$

■

Ejemplo 2. Considere la ED

$$(2xy\,\text{sen}(2x^2y) + 3x^2)\,dx + x^2\,\text{sen}(2x^2y)\,dy = 0$$

Las funciones $M(x, y) = 2xy\,\text{sen}(2x^2y) + 3x^2$, $N(x, y) = x^2\,\text{sen}(2x^2y)$ están definidas en todo el espacio $\mathbb{R}^2$ y son de clase $\mathscr{C}^\infty$ (es decir, de clase $\mathscr{C}^k$ para todo $k \geq 0$). Como

$$\begin{aligned}\frac{\partial M}{\partial y}(x, y) &= \frac{\partial}{\partial y}(2xy\,\text{sen}(2x^2y) + 3x^2) = 4x^3y\cos(2x^2y) + 2x\,\text{sen}(2x^2y) \\ &= \frac{\partial}{\partial x}(x^2\,\text{sen}(2x^2y)) = \frac{\partial N}{\partial x}(x, y)\end{aligned}$$

concluimos que la ecuación es exacta. Para determinar su solución, podemos usar el resultado del apartado (III) del teorema anterior, o bien, proceder paso a paso para obtener ese resultado en este caso particular, como sigue. Debe haber una función $f\colon \mathbb{R}^2 \to \mathbb{R}$ para la que

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= M(x, y) = 2xy\,\text{sen}(2x^2y) + 3x^2 \\ \frac{\partial f}{\partial y}(x, y) &= N(x, y) = x^2\,\text{sen}(2x^2y)\end{aligned}$$

Usando, por ejemplo, la segunda de estas expresiones, tenemos que

$$f(x, y) = \int x^2\,\text{sen}(2x^2y)\,dy + h(x) = -\frac{1}{2}\cos(2x^2y) + h(x)$$

Derivando respecto de x y con $\frac{\partial f}{\partial x}(x, y) = 2xy\,\text{sen}(2x^2y) + 3x^2$ (la primera de las dos expresiones anteriores) obtenemos

$$\frac{\partial}{\partial x}\left(-\frac{1}{2}\cos(2x^2y) + h(x)\right) = 2xy\,\text{sen}(2x^2y) + h'(x) = 2xy\,\text{sen}(2x^2y) + 3x^2$$

de donde $h'(x) = 3x^2$, y entonces $h(x) = x^3$. Concluimos finalmente que la solución de la ecuación dada es la familia de funciones $y = y(x, c)$ dada de forma implícita por $f(x, y) = c$, o sea

$$-\frac{1}{2}\cos(2x^2y) + x^3 = c$$

■

Ejemplo 3. Consideremos la ecuación

$$(x + y^2)\,dx + x(2y\ln x + 1)\,dy = 0$$

Las funciones $M(x, y) = x + y^2$, $N(x, y) = x(2y\ln x + 1)$ son funciones de clase $\mathscr{C}^\infty$ definidas en el abierto simplemente conexo $U = \mathbb{R}^+ \times \mathbb{R} = \{(x, y)|x > 0\} \subseteq \mathbb{R}^2$. Obsérvese que

$$\frac{\partial M}{\partial y}(x, y) = \frac{\partial}{\partial y}(x + y^2) = 2y$$
$$\frac{\partial N}{\partial x}(x, y) = \frac{\partial}{\partial x}(x(2y\ln x + 1)) = 2y + 2y\ln x + 1$$

de modo que en este caso $\frac{\partial M}{\partial y}(x, y) \neq \frac{\partial N}{\partial x}(x, y)$, por lo que la ecuación no es exacta. Nótese sin embargo que si multiplicamos toda la ecuación por $\frac{1}{x}$, ésta se convierte en

$$(1 + \frac{y^2}{x})\,dx + (2y\ln x + 1)\,dy = 0$$

donde ahora

$$\frac{\partial M}{\partial y}(x, y) = \frac{\partial}{\partial y}\left(1 + \frac{y^2}{x}\right) = \frac{2y}{x} = \frac{\partial}{\partial x}(2y\ln x + 1) = \frac{\partial N}{\partial x}(x, y)$$

y la ecuación ya es exacta. Entonces hay una función $f\colon U \to \mathbb{R}$ que

$$\frac{\partial f}{\partial x}(x, y) = M(x, y) = 1 + \frac{y^2}{x}$$
$$\frac{\partial f}{\partial y}(x, y) = N(x, y) = 2y\ln x + 1$$

De la primera de estas expresiones obtenemos que

$$f(x, y) = \int (1 + \frac{y^2}{x})\,dx + g(y) = x + y^2\ln x + g(y)$$

Derivando respecto de y y usando la segunda expresión obtenemos

$$2y\ln x + g'(y) = 2y\ln x + 1$$

de donde $g(y) = y$, y que la solución general de la ecuación es

$$x + y^2\ln x + y = c$$

■

El ejemplo anterior muestra que algunas ecuaciones diferenciales no exactas se pueden convertir en exactas multiplicándolas por un factor adecuado. En general, para la ecuación $M(x, y)\,dx + N(x, y)\,dy = 0$, donde $M, N\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ son funciones de clase $\mathscr{C}^k$, $k \geq 0$, definidas en el abierto U de $\mathbb{R}^2$, se dice que la función no nula $\mu\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$, de clase $\mathscr{C}^k$, es un *factor de integración,* si la ecuación

$$\mu(x, y)M(x, y)\,dx + \mu(x, y)N(x, y)\,dy = 0$$

es exacta. Si $k \geq 1$ y U es simplemente conexo, esto equivale a decir que

$$\frac{\partial}{\partial y}(\mu(x, y)M(x, y)) = \frac{\partial}{\partial x}(\mu(x, y)N(x, y)), \quad \forall(x, y) \in U \tag{F}$$

Por supuesto que nos gustaría "despejar" $\mu(x, y)$ de la expresión anterior, para así poder tener una fórmula general que nos permita hallar factores de integración para cualquier ecuación de primer orden. Sin embargo esto no es posible (si lo fuera, podríamos resolver explícitamente todas las ecuaciones de primer orden y ... —desgraciadamente— esto no es cierto; de hecho son "muy pocas" las ecuaciones de primer orden para las que podemos hacer explícita su solución), por lo que nos debemos conformar con estudiar algunos casos particulares en los que sí podamos hacer tal despeje. Estos casos son:

CASO I. El factor μ depende solamente de la variable X. Digamos que $\mu = \mu(x)$. En este caso la ecuación (F) se ve como

$$\frac{\partial}{\partial y}(\mu(x)M(x, y)) = \frac{\partial}{\partial x}(\mu(x)N(x, y))$$

o bien

$$\mu(x)\frac{\partial M}{\partial y}(x, y) = \mu(x)\frac{\partial N}{\partial x}(x, y) + N(x, y)\mu'(x)$$

de donde

$$\frac{d\mu(x)}{\mu(x)} = \frac{\frac{\partial M}{\partial y}(x, y) - \frac{\partial N}{\partial x}(x, y)}{N(x, y)}\, dx$$

Integrando y despejando $\mu(x)$ se obtiene

$$\mu(x) = \exp\left(\int \frac{\frac{\partial M}{\partial y}(x, y) - \frac{\partial N}{\partial x}(x, y)}{N(x, y)}\, dx\right)$$

Es claro que para que este factor exista es necesario y suficiente que la expresión

$$\frac{\frac{\partial M}{\partial y}(x, y) - \frac{\partial N}{\partial x}(x, y)}{N(x, y)}$$

dependa solamente de la variable x.

Ejemplo 4. Retomemos la ecuación del ejemplo 3

$$(x + y^2)\, dx + x(2y \ln x + 1)\, dy = 0$$

Se tiene que $\frac{\partial M}{\partial y}(x, y) = 2y$, $\frac{\partial N}{\partial x}(x, y) = 2y + 2y \ln x + 1$, y entonces

$$\frac{\frac{\partial M}{\partial y}(x, y) - \frac{\partial N}{\partial x}(x, y)}{N(x, y)} = \frac{2y - 2y - 2y \ln x - 1}{x(2y \ln x + 1)} = -\frac{1}{x}$$

de modo que la ecuación posee un factor de integración que depende solamente de x, el cual es

$$\mu(x) = \exp\left(-\int \frac{1}{x}dx\right) = \frac{1}{x}$$

tal como se aplicó en el ejemplo 3. ■

Ejemplo 5. Un tipo muy importante de ecuaciones de primer orden lo constituyen las *ecuaciones lineales*, que son ecuaciones de la forma

$$y' + P(x)y = Q(x)$$

donde P y Q son funciones continuas de X. Hay varias formas de abordar el problema de la solución de una ecuación lineal. En este ejemplo veremos cómo resolverla con las ideas de ecuaciones exactas desarrolladas en esta sección (en los ejercicios al final de la misma se verán algunas alternativas más de solución). Escribamos entonces la ecuación $y' + P(x)y = Q(x)$ como

$$(Q(x) - P(x)y)\,dx - dy = 0$$

Siendo $M(x, y) = Q(x) - P(x)y$, $N(x, y) = -1$, $\frac{\partial M}{\partial y}(x, y) = -P(x)$, $\frac{\partial N}{\partial x}(x, y) = 0$, es claro que la ecuación no es exacta. Sin embargo, obsérvese que

$$\frac{\dfrac{\partial M}{\partial y}(x, y) - \dfrac{\partial N}{\partial x}(x, y)}{N(x, y)} = P(x)$$

por lo que es posible obtener un factor de integración que dependa solamente de x, a saber $\mu = \mu(x) = e^{\int P(x)\,dx}$. Multiplicando toda la ecuación por este factor se obtiene

$$e^{\int P(x)\,dx}(Q(x) - P(x)y)\,dx - e^{\int P(x)\,dx}\,dy = 0$$

donde ahora $M(x, y) = e^{\int P(x)\,dx}(Q(x) - P(x)y)$, $N(x, y) = -e^{\int P(x)dx}$, $\frac{\partial M}{\partial y}(x, y) = -P(x)e^{\int P(x)\,dx} = \frac{\partial N}{\partial x}(x, y)$; es decir, la ecuación es exacta. Hay entonces una función $f(x, y)$ tal que

$$\frac{\partial f}{\partial x}(x, y) = e^{\int P(x)dx}(Q(x) - P(x)y)$$
$$\frac{\partial f}{\partial y}(x, y) = -e^{\int P(x)dx}$$

De la segunda de estas expresiones obtenemos

$$f(x, y) = \int -e^{\int P(x)\,dx}\,dy + h(x) = -e^{\int P(x)\,dx}y + h(x)$$

Derivando respecto de x, e imponiendo la correspondiente condición sobre esta derivada parcial, se tiene

$$\frac{\partial f}{\partial x}(x, y) = -P(x)e^{\int P(x)\,dx}y + h'(x) = e^{\int P(x)\,dx}(Q(x) - P(x)y)$$

de donde

$$h(x) = \int Q(x)e^{\int P(x)\,dx}\,dx$$

y entonces, la solución general de la ecuación lineal es

$$f(x, y) = -e^{\int P(x)\,dx}y + \int Q(x)e^{\int P(x)\,dx}\,dx = c$$

o sea

$$y = e^{-\int P(x)\,dx}\int Q(x)e^{\int P(x)\,dx}\,dx + ce^{-\int P(x)\,dx}$$ ■

CASO II. El factor μ depende sólo de la variable y. Digamos que $\mu = \mu(y)$. En este caso la ecuación (F) se ve como

$$\frac{\partial}{\partial y}(\mu(y)M(x, y)) = \frac{\partial}{\partial x}(\mu(y)N(x, y))$$

o bien

$$\mu(y)\frac{\partial M}{\partial y}(x, y) + M(x, y)\mu'(y) = \mu(y)\frac{\partial N}{\partial x}(x, y)$$

de donde

$$\frac{d\mu(y)}{\mu(y)} = \frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)}dy$$

Integrando y despejando $\mu(y)$ se obtiene

$$\mu(y) = \exp\left(\int \frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)}dy\right)$$

Es claro que para que este factor exista es necesario y suficiente que la expresión

$$\frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)}$$

dependa solamente de la variable y.

Ejemplo 6. Consideremos la ecuación

$$y(5x^4 + y)dx + x[x^4(5y + 2) + y(5y + 3)]dy = 0$$

Las funciones $M(x, y) = y(5x^4 + y)$, $N(x, y) = x[x^4(5y + 2) + y(5y + 3)]$ son funciones de clase $\mathscr{C}^\infty$ definidas en todo $\mathbb{R}^2$. Se tiene

$$\frac{\partial M}{\partial y}(x, y) = 5x^4 + 2y$$

$$\frac{\partial N}{\partial x}(x, y) = 25x^4y + 10x^4 + 5y^2 + 3y$$

por lo que la ecuación no es exacta. Por otra parte se tiene

$$\frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)} = \frac{25x^4y + 10x^4 + 5y^2 + 3y - 5x^4 - 2y}{y(5x^4 + y)}$$
$$= 5 + \frac{1}{y}$$

así la ecuación posee un factor de integración que depende sólo de la variable y. Este es

$$\mu(y) = \exp\left(\int (5 + \frac{1}{y})dy\right) = ye^{5y}$$

Multiplicando toda la ecuación diferencial dada por este factor obtenemos

$$y^2e^{5y}(5x^4 + y)\,dx + xye^{5y}[x^4(5y + 2) + y(5y + 3)]\,dy = 0$$

Se tiene ahora

$$\frac{\partial M}{\partial y}(x, y) = ye^{5y}(25x^4y + 10x^4 + 5y^2 + 3y) = \frac{\partial N}{\partial x}(x, y)$$

y ahora la ecuación es exacta. Debe haber entonces una función $f(x, y)$ tal que

$$\frac{\partial f}{\partial x}(x, y) = y^2e^{5y}(5x^4 + y)$$
$$\frac{\partial f}{\partial y}(x, y) = xye^{5y}[x^4(5y + 2) + y(5y + 3)]$$

Es claro que para determinar la función $f(x, y)$ resulta más conveniente comenzar con la primera de estas expresiones. De ella se obtiene

$$f(x, y) = \int \left(y^2e^{5y}(5x^4+)\right) dx + g(y) = xy^2e^{5y}(x^4 + y) + g(y)$$

Derivando respecto de y y usando la segunda expresión se obtiene

$$\frac{\partial f}{\partial y}(x, y) = xye^{5y}[x^4(5y + 2) + y(5y + 3)] + g'(y)$$
$$= xye^{5y}[x^4(5y + 2) + y(5y + 3)]$$

de donde $g(y)$ es constante, y así la solución general de la ecuación es

$$xy^2e^{5y}(x^4 + y) = c$$

■

CASO III. El factor μ depende de la variable z donde $z = z(x, y)$. Digamos que $\mu = \mu(z)$, donde z es una función de las variables x, y. En este caso la ecuación (F) se ve como

$$\frac{\partial}{\partial y}(\mu(z)M(x, y)) = \frac{\partial}{\partial x}(\mu(z)N(x, y))$$

o bien

$$\mu(z)\frac{\partial M}{\partial y}(x, y) + M(x, y)\mu'(z)\frac{\partial z}{\partial y}(x, y) = \mu(z)\frac{\partial N}{\partial x}(x, y) + N(x, y)\mu'(z)\frac{\partial z}{\partial x}(x, y)$$

de donde

$$\frac{d\mu(z)}{\mu(z)} = \frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)\frac{\partial z}{\partial y}(x, y) - N(x, y)\frac{\partial z}{\partial x}(x, y)}\, dz$$

Integrando y despejando $\mu(z)$ se obtiene

$$\mu(z) = \exp\left(\int \frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)\frac{\partial z}{\partial y}(x, y) - N(x, y)\frac{\partial z}{\partial x}(x, y)}\, dz\right)$$

Es claro que para que este factor se dé es necesario y suficiente que la expresión

$$\frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)\frac{\partial z}{\partial y}(x, y) - N(x, y)\frac{\partial z}{\partial x}(x, y)}\, dz$$

dependa solamente de la variable z.

Ejemplo 7. Consideremos la ecuación

$$(2x^2y + 6xy + 1)\, dx + x^2(x + 3)\, dy = 0$$

Se tiene

$$\frac{\partial M}{\partial y}(x, y) = 2x^2 + 6x, \qquad \frac{\partial N}{\partial x}(x, y) = 3x^2 + 6x$$

de donde se ve que la ecuación no es exacta. Procuremos un factor de integración que dependa de la variable $z = x^2y$ (esta información se debe proporcionar en el ejercicio). Se tiene

$$\frac{\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)}{M(x, y)\frac{\partial z}{\partial y}(x, y) - N(x, y)\frac{\partial z}{\partial x}(x, y)}\, dz = \frac{3x^2 + 6x - 2x^2 - 6x}{(2x^2y + 6xy + 1)x^2 - x^2(x + 3)2xy} = 1$$

de modo que la ecuación posee un factor de integración que depende solamente de z. Este es

$$\mu(z) = \exp\left(\int dz\right) = e^z = e^{x^2y}$$

Multiplicando toda la ecuación por este factor se obtiene

$$e^{x^2y}(2x^2y + 6xy + 1)dx + e^{x^2y}x^2(x + 3)dy = 0$$

Ahora se tiene

$$\frac{\partial M}{\partial y}(x, y) = xe^{x^2 y}(2x^3 y + 6x^2 y + 3x + 6) = \frac{\partial N}{\partial x}(x, y)$$

por lo que ahora la ecuación es exacta. Debe haber entonces una función $f(x, y)$ con

$$\frac{\partial f}{\partial x}(x, y) = e^{x^2 y}(2x^2 y + 6xy + 1)$$
$$\frac{\partial f}{\partial y}(x, y) = e^{x^2 y}x^2(x + 3)$$

De la segunda de estas expresiones se obtiene que

$$f(x, y) = \int e^{x^2 y}x^2(x + 3)\,dy + h(x) = e^{x^2 y}(x + 3) + h(x)$$

Derivando respecto de x, y usando la segunda expresión se obtiene

$$\frac{\partial f}{\partial x}(x, y) = e^{x^2 y}(2x^2 y + 6xy + 1) + h'(x) = e^{x^2 y}(2x^2 y + 6xy + 1)$$

de donde $h(x)$ es una constante y la solución general de la ecuación es

$$e^{x^2 y}(x + 3) = c$$

■

Ejercicios (Capítulo 7, Sección 6)

Verifique en los ejercicios 1–16 que la ecuación dada es exacta y obtenga su solución general.

1. $(2xy + 1)\,dx + (x^2 + 4y)\,dy = 0$

2. $(y + 1)\,dx + (x + 1)\,dy = 0$

3. $(6x^5 y + 7)\,dx + (x^6 + 9)\,dy = 0$

4. $(2xy^2 + y + 5)\,dx + (2x^2 y + x + 2)\,dy = 0$

5. $(4xy^3 + 5y^4 + 4)\,dx + 2(3x^2 y^2 + 10xy^3 + 4)\,dy = 0$

6. $(2xy^3 + y + 1)\,dx + (3x^2 y^2 + x + 7)\,dy = 0$

7. $(3y^3 + y + 2)\,dx + x(9y^2 + 1)\,dy = 0$

8. $x(12y^3 + 1)\,dx + y(18x^2 y + 1)\,dy = 0$

9. $2x(y^3 + 1)\,dx + 3y^2(x^2 + 1)\,dy = 0$

10. $(4x + \operatorname{sen}^2 y)\,dx + (x \operatorname{sen} 2y + 1)\,dy = 0$

11. $y(6y^6 + 5)\,dx + [x(42y^6 + 5) + 1]\,dy = 0$

12. $2(y^2 + 5y + 1)\,dx + [2x(2y + 5) + 3]\,dy = 0$

13. $(3x^2 + 10xy + 10y)\,dx + (5x^2 + 10x + y)\,dy = 0$

14. $(y^2 e^{x+y} + 1)\,dx + [ye^{x+y}(y + 2) + 1]\,dy = 0$

15. $(x \cos(x + y) + 1)dy + (x \cos(x + y) + \operatorname{sen}(x + y) + 1)dx = 0$

16. $[y \cos(x + y) + \cos(x - y) - x \operatorname{sen}(x - y)]\,dx + [y \cos(x + y) + \operatorname{sen}(x + y) + x \operatorname{sen}(x - y)]\,dy = 0$

En los ejercicios 17–23, obtenga un factor de integración para la ecuación dada. Halle entonces su solución general.

17. $y(x - 2y + 1)\,dx + (x - 4y)\,dy = 0$

18. $y(3x^2 + y)\,dx + x(2x^2 + 3y)\,dy = 0$

19. $y[15x + 4y(y + 12)]\,dx + 3x[x + y(y + 8)]\,dy = 0$

20. $y[9x + 2(7y + 6)]\,dx + 3x(2x + 7y + 4)\,dy = 0$. (Sugerencia: procúrese un factor de integración del tipo $\mu = \mu(xy)$).

21. $2x(6x^2 + y + 6)\,dx + (12x^2y + 2y^2 + 1)\,dy = 0$. (Sugerencia: procúrese un factor de integración del tipo $\mu = \mu(x^2 + y^2)$).

22. $5x^4(3y + 1)\,dx + [x^5 + y^4(16y + 5)]\,dy = 0$. (Sugerencia: procúrese un factor de integración del tipo $\mu = \mu(x^5 + y^5)$).

23. $[3(x + y^2)\ln(x + y^2) + 3x + 4y]\,dx + 2[2(x + y^2)\ln(x + y^2) + y(3x + 4y)]\,dy = 0$. (Sugerencia: procúrese un factor de integración del tipo $\mu = \mu(x + y^2)$).

En los ejercicios 24–28 se dan ecuaciones lineales de primer orden. Siga el procedimiento descrito en el ejemplo 5 para obtener la solución general de estas ecuaciones.

24. $y' + y = \cos x$

25. $y' + 2y = e^{3x}$

26. $y' + y/x = x$

27. $y' + \dfrac{x}{x^2 + 1}y = 1$

28. $y' + 3y = \operatorname{sen}^2 x$

29. (Más sobre ecuaciones lineales). Un procedimiento de alternativa al considerado en el ejemplo 5 para obtener la solución general de la ecuación lineal de primer orden $y' + P(x)y = Q(x)$ es como sigue: multiplique ambos miembros de la ecuación por el factor $e^{\int P(x)\,dx}$, para obtener $e^{\int P(x)\,dx}y' + P(x)e^{\int P(x)\,dx}y = Q(x)e^{\int P(x)\,dx}$. Compruebe que el primer miembro de esta ecuación es la derivada (respecto de x) del producto $e^{\int P(x)\,dx}y$. Concluya entonces que este producto es igual $a \int Q(x)e^{\int P(x)\,dx}\,dx + c$, donde c es una constante arbitraria. Despeje la función y para obtener la solución general de la ecuación como

$$y = e^{-\int P(x)\,dx}\int Q(x)e^{\int P(x)\,dx}\,dx + ce^{-\int P(x)\,dx}$$

Siga este procedimiento para reobtener la solución general de las ecuaciones lineales de los ejercicios 31–35.

30. (Ecuaciones de Bernoulli). Una ecuación del tipo $y' + P(x)y = Q(x)y^n$, en la que $n \neq 0, 1$, se llama ecuación de Bernoulli. Demuestre que si se cambia la función $y = y(x)$ por la función $z = z(x)$, según la fórmula $z = y^{1-n}$, dicha ecuación se reduce a la ecuación lineal $z' + (1 - n)P(x)z = (1 - n)Q(x)$. Obtenga la solución general de las siguientes ecuaciones de Bernoulli: **a.** $y' + y = y^2$; **b.** $y' + 3y = e^x y^2$; **c,** $y' + xy = xy^3$.

31. (Ecuaciones homogéneas). Se dice que una ecuación diferencial de primer orden del tipo $y' = f(x, y)$ es homogénea si la función $f(x, y)$ es homogénea de grado cero.

a. Considere la siguiente definición de ecuación homogénea: la ecuación diferencial $M(x, y)\,dx + N(x, y)\,dy = 0$ se dice ser homogénea si las funciones $M(x, y)$ y $N(x, y)$ son homogéneas del mismo grado. Demuestre que esta definición es equivalente a la dada al inicio del ejercicio.

b. Demuestre que si la función $f(x, y)$ es homogénea de grado cero, ésta se puede escribir como $f(x, y) = \varphi(y/x)$, para una función φ de la variable $u = y/x$.

c. Considere la ecuación homogénea $y' = f(x, y)$. Escribiendo ésta como $y' = \varphi(y/x)$, demuestre que cambiando la función $y = y(x)$ por la función $u = u(x)$, según la fórmula $y = ux$, la ecuación puede ser escrita como $xu' = \varphi(u) - u$, la cual se integra directamente como $\int \frac{du}{\varphi(u)-u} = \int \frac{dx}{x} + c$.

d. Siga el procedimiento descrito en el inciso anterior para resolver las ecuaciones siguientes:

d1. $(-8x + 17y)\,dx - 3(4y - x)\,dy = 0$

d2. $(6x + y)\,dx + (x + 7y)\,dy = 0$

d3 $(x + (x^2 + y^2)^{1/2})\,dx + y\,dy = 0$

d4 $(x + y)\,dx - 2y\,dy = 0$

32. (Ecuaciones homogéneas desde el punto de vista de ecuaciones exactas). Considere la ecuación homogénea $M(x, y)\,dx + N(x, y)\,dy = 0$. Esta ecuación no es, en general, exacta. Demuestre que $\mu(x, y) = \frac{1}{xM(x,y)+yN(x,y)}$ es un factor de integración para la ecuación. (Sugerencia: para demostrar que la ecuación $\mu(x, y)M(x, y)\,dx + \mu(x, y)N(x, y)\,dy = 0$ es exacta, i.e. para demostrar la igualdad de las derivadas parciales $\frac{\partial}{\partial y}(\mu(x, y)M(x, y)) = \frac{\partial}{\partial x}(\mu(x, y)N(x, y))$, aplique el teorema de Euler sobre funciones homogéneas presentado en el capítulo 2).

33. Demuestre que si la ecuación $M(x, y)\,dx + N(x, y)\,dy = 0$ es homogénea y exacta, entonces $xM(x, y) + yN(x, y) = c$ es su solución general.

7.7 Integrales de línea respecto a la longitud de arco

En esta sección estudiaremos un tipo distinto de integrales de línea de las que ya hemos presentado en la sección 3, donde se estudiaron integrales de *campos vectoriales* a lo largo de caminos en el espacio. Ahora estudiaremos integrales sobre caminos (las regiones de integración serán las mismas) de *funciones reales definidas en (algún subconjunto de)* $\mathbb{R}^n$. Dividiremos la presente sección en dos subsecciones, en la primera de las cuales presentamos la definición y las propiedades de estas integrales y en la segunda algunas de sus aplicaciones.

7.7.1 Definición y propiedades

Antes de establecer formalmente la definición de estas nuevas integrales, consideremos un par de ejemplos preliminares en cuya solución aparecerán de manera natural las expresiones que, veremos, serán las nuevas integrales objeto de estudio en esta sección.

Ejemplo 1. (La masa total de un alambre) Pensemos en un alambre ℓ, en el plano o en el espacio, del cual conocemos su densidad lineal (en gr/cm) en cada punto, por la función $\rho\colon \ell \to \mathbb{R}$ que asocia a cada punto $\mathbf{p} \in \ell$, el número real $\rho(\mathbf{p})$ = densidad lineal del alambre en $\mathbf{p}$. Supongamos que el alambre ℓ es la imagen de un camino (regular, de clase $\mathscr{C}^1$, simple) $\boldsymbol{\lambda}\colon [a, b] \to \mathbb{R}^2$ o $\mathbb{R}^3$. Se quiere entonces calcular la masa total de ℓ. Observe que si la densidad del alambre fuera constante, digamos ρ_0 gr/cm, el problema se resulve inmediatamente multiplicando ρ_0 por la longitud total del alambre (escribámosla como L_ℓ), la cual, en términos del camino $\boldsymbol{\lambda}$ es, como sabemos (sección 5 capítulo 5)

$$L_\ell = \int_a^b \|\boldsymbol{\lambda}'(t)\|\, dt$$

Es decir, en este caso la masa total del alambre M_ℓ es

$$M_\ell = \rho_0 L_\ell = \rho_0 \int_a^b \|\boldsymbol{\lambda}'(t)\|\, dt$$

Vayamos ahora al caso general: la densidad lineal del alambre es variable. Tomemos una partición $\mathbf{p}$ del intervalo $[a, b]$ como sigue

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b$$

A esta partición corresponde una partición en ℓ, digamos

$$\mathbf{p}_0 < \mathbf{p}_1 < \mathbf{p}_2 < \cdots < \mathbf{p}_{n-1} < \mathbf{p}_n$$

en donde $\mathbf{p}_i = \boldsymbol{\lambda}(t_i) \in \ell$, $i = 0, 1, 2, \ldots, n$. En cada subintervalo $[t_{i-1}, t_i]$ de la partición de $[a, b]$, tomemos un punto arbitrario ξ_i. A este punto $\xi_i \in [t_{i-1}, t_i]$ corresponde el punto $\zeta_i = \boldsymbol{\lambda}(\xi_i) \in \ell$ el cual se encuentra entre $\mathbf{p}_{i-1}$ y $\mathbf{p}_i$.

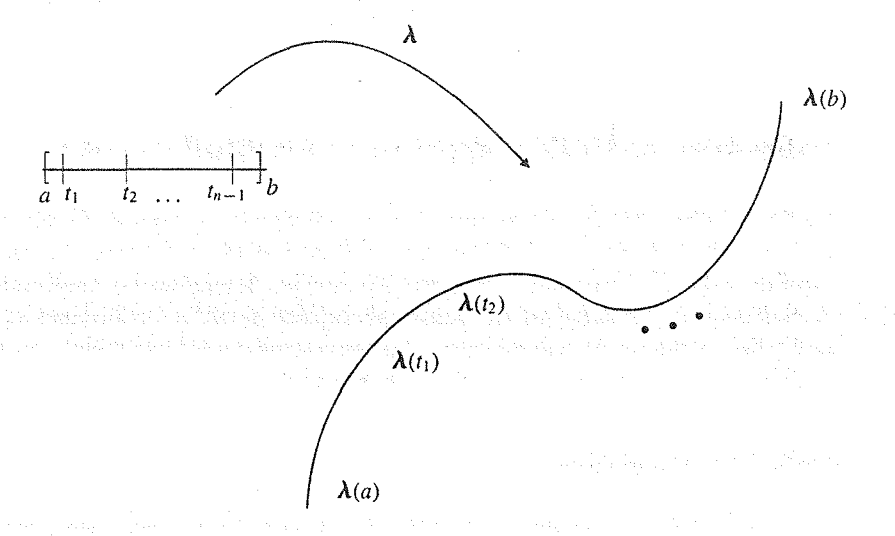


Figura 1. La partición P de $[a, b]$ y la correspondiente partición de ℓ.

La masa del alambre entre los puntos $\mathbf{p}_{i-1}$ y $\mathbf{p}_i$ es aproximadamente igual a la densidad del alambre en ζ_i multiplicada por la longitud del alambre entre esos dos puntos. Es decir

$$M_i = \text{Masa del alambre entre } \mathbf{p}_{i-1} \text{ y } p_i \approx \rho(\zeta_i) \int_{t_{i-1}}^{t_i} \|\boldsymbol{\lambda}'(t)\|\, dt$$

Siguiendo el teorema del valor medio para integrales podemos escribir que

$$M_i \approx \rho(\zeta_i) \int_{t_{i-1}}^{t_i} \|\boldsymbol{\lambda}'(t)\|\, dt \approx \rho(\zeta_i)\|\boldsymbol{\lambda}'(\xi_i)\|(t_i - t_{i-1})$$

de modo que la masa total del alambre se puede escribir como (con $\zeta_i = \boldsymbol{\lambda}(\xi_i)$)

$$M_\ell \approx \sum_{i=1}^{n} M_i \approx \sum_{i=1}^{n} \rho(\boldsymbol{\lambda}(\xi_i))\|\boldsymbol{\lambda}'(\xi_i)\|(t_i - t_{i-1})$$

Esta es una expresión típica de sumas de Riemann para la función $\phi\colon [a, b] \to \mathbb{R}$, $\phi(t) = \rho(\boldsymbol{\lambda}(t))$ $\|\boldsymbol{\lambda}'(t)\|$, cuyo paso al límite cuando n tiende a infinito y la norma de la partición p tiende a cero, nos da el valor exacto de la masa que estamos procurando. O sea

$$\begin{aligned} M_\ell &= \lim_{\substack{n\to\infty \\ \text{norma de } P\to 0}} \sum_{i=1}^{n} \rho(\boldsymbol{\lambda}(\xi_i))\|\boldsymbol{\lambda}'(\xi_i)\|(t_i - t_{i-1}) \\ &= \int_a^b \rho(\boldsymbol{\lambda}(t))\|\boldsymbol{\lambda}'(t)\|\, dt \end{aligned}$$

■

Ejemplo 2. (El pez come algas) En un estanque (veámoslo como un subconjunto de $\mathbb{R}^3$) lleno de algas se deposita un pez π que se alimenta de ellas. Cuando el pez va a comer siempre comienza su nado en un punto determinado del estanque y sigue una trayectoria ℓ que se puede considerar como la imagen de un camino (de clase $\mathscr{C}^1$) $\boldsymbol{\lambda}\colon [a, b] \to \mathbb{R}^3$. Conociendo la densidad lineal (gr/cm) de algas sobre la trayectoria del pez, digamos que dada por la función $\rho\colon \ell \to \mathbb{R}$, se quiere saber la cantidad total de algas que el pez come en su recorrido. Nuevamente se puede ver que si la densidad de algas en el estanque fuera constante, o, al menos fuera constante sobre la trayectoria que el pez sigue, entonces la cantidad total de algas que el pez comerá se obtiene como el producto de la densidad constante por la longitud de la curva que el pez recorre. En caso contrario, un argumento análogo al presentado en el ejemplo anterior (dejamos al lector copiar los detalles) nos conduce a que la cantidad total de algas que el pez come está dada por

$$\int_a^b \rho(\boldsymbol{\lambda}(t))\|\boldsymbol{\lambda}'(t)\|\, dt$$

■

La expresión obtenida en los dos ejemplos anteriores es lo que a continuación definiremos como una integral de línea con respecto a la longitud de arco.

Definición. Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función real continua definida en el abierto U de $\mathbb{R}^n$ y sea $\boldsymbol{\lambda}\colon [a, b] \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$ de modo que $\boldsymbol{\lambda}([a, b]) \subset U$. Se define la *integral*

de línea con respecto a la longitud de arco de la función f a lo largo del camino $\boldsymbol{\lambda}$ (o sobre el camino $\boldsymbol{\lambda}$), denotada por $\int_{\boldsymbol{\lambda}} f\,ds$, como

$$\int_{\boldsymbol{\lambda}} f\,ds = \int_a^b f(\boldsymbol{\lambda}(t))\|\boldsymbol{\lambda}'(t)\|\,dt$$ ■

Nótese que las hipótesis de continuidad de la función f y de la clase de diferenciabilidad del camino $\boldsymbol{\lambda}$ en la definición anterior, garantizan la existencia de la integral que aparece en ella. En efecto, en tal caso la función $\phi(t) = f(\boldsymbol{\lambda}(t))\|\boldsymbol{\lambda}'(t)\|$ es continua en $[a, b]$ y por lo tanto es integrable. La definición anterior sirve para caminos $\boldsymbol{\lambda}$ que son seccionalmente $\mathscr{C}^1$. En este caso podemos escribir $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \cdots + \boldsymbol{\lambda}_k$, donde cada uno de los caminos $\boldsymbol{\lambda}_i$ es de clase $\mathscr{C}^1$ y entonces

$$\int_{\boldsymbol{\lambda}} f\,ds = \sum_{i=1}^{n} \int_{\boldsymbol{\lambda}_i} f\,ds$$

La terminología usada para designar este nuevo tipo de integrales de línea se puede justificar viendo la relación entre estas nuevas integrales de línea y las integrales de línea que hemos venido estudiando en este capítulo desde la sección 3. En efecto, considere un campo continuo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ definido en el abierto U de $\mathbb{R}^n$. Sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino regular de clase $\mathscr{C}^1$ cuya imagen está contenida en U. La integral de línea del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}$ es, como sabemos

$$\begin{aligned} \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\,dt \\ &= \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \frac{\boldsymbol{\lambda}'(t)}{\|\boldsymbol{\lambda}'(t)\|}\|\boldsymbol{\lambda}'(t)\|\,dt = \int_{\boldsymbol{\lambda}} f\,ds \end{aligned}$$

donde $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es tal que $f(\boldsymbol{\lambda}(t)) = \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \frac{\boldsymbol{\lambda}'(t)}{\|\boldsymbol{\lambda}'(t)\|}$. Nótese que los valores de la función f en los puntos $\boldsymbol{\lambda}(t)$ de la traza del camino $\boldsymbol{\lambda}$ son entonces las *proyecciones* del campo $\mathbf{F}$ sobre el vector velocidad $\boldsymbol{\lambda}'(t)$ normalizado. Más aún, son las *componentes ortogonales* de los vectores $\mathbf{F}$ sobre los vectores velocidad normalizados del camino $\boldsymbol{\lambda}$ (ver sección 3 del capítulo 1). Denotando como $\mathbf{T}(t)$ al vector tangente unitario a la curva $\boldsymbol{\lambda}$, podemos escribir en forma abreviada a f como $f = \mathbf{F} \cdot \mathbf{T}$ (o sea que $f(\boldsymbol{\lambda}(t)) = \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \mathbf{T}(t)$). Estas componentes son las que se integran *con respecto de la longitud del camino*, pues, recordando que para el camino $\boldsymbol{\lambda}$ se tiene definida su longitud del punto inicial $\boldsymbol{\lambda}(a)$ a un punto $\boldsymbol{\lambda}(t)$ cualquiera como

$$s = s(t) = \int_a^t \|\boldsymbol{\lambda}'(u)\|\,du$$

se tiene, por el Teorema Fundamental del Cálculo, que $ds = \|\boldsymbol{\lambda}'(t)\|\,dt$. Así pues, el símbolo ds que aparece en la notación de las integrales de línea de la función f, es de hecho la diferencial de la longitud de arco del camino $\boldsymbol{\lambda}$, siendo entonces esta longitud de arco la "variable de integración" en la integral $\int_{\boldsymbol{\lambda}} f\,ds$. En resumen, la integral del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$, $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$, es la misma que la integral de línea respecto de la longitud de arco $\int_{\boldsymbol{\lambda}} f\,ds = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot T\,ds$ de la función f que asocia en cada punto $\boldsymbol{\lambda}(t)$ del camino $\boldsymbol{\lambda}$ la componente ortogonal $\mathbf{F} \cdot \mathbf{T}$ del campo $\mathbf{F}$ sobre el vector velocidad $\boldsymbol{\lambda}'(t)$ normalizado (el vector tangente unitario $T(t)$).

Por último, nótese que si tomamos la función constante $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = 1$, entonces

$$\int_{\boldsymbol{\lambda}} f\,ds = \int_{\boldsymbol{\lambda}} ds = \int_a^b \|\boldsymbol{\lambda}'(t)\|\,dt = \text{longitud del camino } \boldsymbol{\lambda}.$$

Compare esto con los hechos ya considerados en el estudio de las integrales dobles y triples, donde si se integraba la función constante 1, el resultado era (numéricamente) igual a la "medida" de la región de integración J, es decir, al área de $J \subset \mathbb{R}^2$ en el caso de integrales dobles, o al volumen de $J \subset \mathbb{R}^3$ en el caso de integrales triples. En nuestro caso se tiene una situación completamente análoga: al integrar (al calcular la integral de línea respecto de la longitud de arco de) la función $f(x) = 1$ se obtiene la medida de la región de integración, que (en este caso) es la longitud del camino $\boldsymbol{\lambda}$ sobre el que estamos integrando.

Veamos algunos ejemplos.

Ejemplo 3. Calculemos la integral de línea de la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$, a lo largo del camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\lambda}} xy\,ds &= \int_0^1 (t)(t^2)\|(1, 2t)\|\,dt = \int_0^1 t^3\sqrt{1+4t^2}\,dt \\ &\overset{1+4t^2=u^2}{=} \frac{1}{16}\int_1^{\sqrt{5}} u^2(u^2-1)du = \frac{1}{16}\left[\frac{1}{5}u^5 - \frac{1}{3}u^3\right]_1^{\sqrt{5}} \\ &= \frac{5\sqrt{5}}{24} + \frac{1}{120}\end{aligned}$$

■

Ejemplo 4. Calculemos la integral de línea de la misma función $f(x, y) = xy$ del ejemplo anterior, sobre el camino $\boldsymbol{\mu}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\mu}(t) = (1-t, (1-t)^2)$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\mu}} xy\,ds &= \int_0^1 (1-t)(1-t)^2\sqrt{1+4(1-t)^2}\,dt \overset{1-t=T}{=} \\ &= \int_0^1 T^3\sqrt{1+4T^2}\,dT = \frac{5\sqrt{5}}{24} + \frac{1}{120}\end{aligned}$$

Nótese que el camino $\boldsymbol{\mu}$ considerado en este ejemplo es el negativo del camino $\boldsymbol{\lambda}$ del ejemplo anterior, es decir, $\boldsymbol{\mu}$ recorre la misma imagen que $\boldsymbol{\lambda}$ pero en sentido inverso. Así, con este par de ejemplos podemos concluir que

$$\int_{\boldsymbol{\lambda}} xy\,ds = \int_{\boldsymbol{\mu}} xy\,ds = \int_{-\boldsymbol{\lambda}} xy\,ds$$

Esta es una propiedad general de estas nuevas integrales de línea que estudiaremos más adelante: estas integrales no cambian de signo (como lo hacen las integrales de línea de campos vectoriales) al cambiar la orientación del camino sobre el que se está integrando. Sugerimos al lector que desde este momento piense esta propiedad en función de los dos ejemplos preliminares presentados en esta subsección: es claro que la masa total del alambre (en el ejemplo 1) será la misma si la calculamos "de ida" que "de regreso", o bien, que el "pez come algas" (ejemplo 2) come la misma cantidad de algas recorriendo su trayectoria en un sentido que en el inverso. ■

Ejemplo 5. Calculemos la integral de línea de la función $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2$ sobre el camino $\boldsymbol{\lambda}: [-\pi, \pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (-\operatorname{sen}(t/2), \cos(t/2))$. Obsérvese que este camino recorre una vez el semicírculo superior $x^2 + y^2 = 1$ en sentido antihorario. Como $f(\boldsymbol{\lambda}(t)) = 1$, se tiene

$$\int_{\boldsymbol{\lambda}} f\, ds = \int_{-\pi}^{\pi} \|\boldsymbol{\lambda}'(t)\| = \int_{-\pi}^{\pi} \|((-1/2)\cos(t/2), (-1/2)\operatorname{sen}(t/2))\|\, dt$$

$$= \frac{1}{2}\int_{-\pi}^{\pi} dt = \pi$$

■

Ejemplo 6. Con la misma función f del ejemplo anterior, calculemos su integral de línea sobre el camino $\boldsymbol{\mu}: [-\pi, \pi] \to \mathbb{R}^2$, $\boldsymbol{\mu}(t) = (-\cos(t^2/\pi), \operatorname{sen}(t^2/\pi))$. Nótese que este camino tiene la misma traza que el camino $\boldsymbol{\lambda}$ del ejemplo anterior, pero ésta es recorrida *dos veces*: va del punto $\boldsymbol{\mu}(-\pi) = (1, 0)$ al punto $\boldsymbol{\mu}(0) = (-1, 0)$ por el semicírculo unitario superior, y luego regresa al punto $\boldsymbol{\mu}(\pi) = (1, 0)$. Nuevamente se tiene que $f(\boldsymbol{\mu}(t)) = 1$. Entonces

$$\int_{\boldsymbol{\mu}} f\, ds = \int_{-\pi}^{\pi} \|\boldsymbol{\mu}'(t)\|\, dt$$

$$= \int_{-\pi}^{\pi} \|(2t/\pi)\operatorname{sen}(t^2/\pi), (2t/\pi)\cos(t^2/\pi))\|\, dt$$

$$= \int_{-\pi}^{\pi} |2t/\pi|\, dt = \frac{2}{\pi}\int_{-\pi}^{\pi} |t|\, dt = \frac{4}{\pi}\int_{0}^{\pi} t\, dt = 2\pi$$

Obsérvese que, recorriendo el camino $\boldsymbol{\mu}$ el doble del recorrido del camino $\boldsymbol{\lambda}$ (sobre la misma traza), el resultado que obtuvimos es *el doble* del correspondiente del ejemplo anterior. Esta será también otra propiedad importante de las integrales de línea, respecto de la longitud de arco, que contrasta con la correspondiente propiedad de las integrales de línea estudiadas anteriormente. Esta se refiere al carácter *acumulativo* de este tipo de integrales. Piense en el ejemplo 2: si el pez recorre dos veces la misma trayectoria, es claro que comerá el doble de algas. ■

Al igual que las integrales de línea de campos vectoriales, estas nuevas integrales tienen la propiedad de linealidad: si $f, g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ son dos funciones continuas definidas en el abierto U de $\mathbb{R}^n$ y $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ es un camino seccionalmente $\mathscr{C}^1$, entonces

$$\int_{\boldsymbol{\lambda}} (f + kg)\, ds = \int_{\boldsymbol{\lambda}} f\, ds + k\int_{\boldsymbol{\lambda}} g\, ds$$

donde $k \in \mathbb{R}$. Dejamos al lector que presente una prueba de este hecho.

En el teorema siguiente se establece la importante propiedad de las integrales de línea con respecto a la longitud de arco que ya ha sido anunciada en el ejemplo 4.

Teorema 7.7.1 Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función continua definida en el abierto U de $\mathbb{R}^n$. Sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$, de modo que $\boldsymbol{\lambda}([a, b]) \subset U$ y sea $\boldsymbol{\mu}: [c, d] \to \mathbb{R}^n$ una reparametrización de $\boldsymbol{\lambda}$, digamos $\boldsymbol{\mu} = \boldsymbol{\lambda} \circ \phi$, donde $\phi: [c, d] \to [a, b]$ es una función sobreyectiva de clase $\mathscr{C}^1$ tal que $\phi'(t) \neq 0\ \forall t \in [c, d]$. Entonces

$$\int_{\boldsymbol{\lambda}} f\, ds = \int_{\boldsymbol{\mu}} f\, ds$$

Demostración. Se tiene

$$\int_{\mu} f\, ds = \int_c^d f(\mu(t))\|\mu'(t)\|\, dt = \int_c^d f(\lambda(\phi(t)))\|\lambda'(\phi(t))\phi'(t)\|\, dt$$

$$= \int_c^d f(\lambda(\phi(t)))\|\lambda'(\phi(t))\||\phi'(t)|\, dt = \longrightarrow \qquad (*)$$

Si $\phi'(t) > 0\ \forall t \in [c, d]$, se tiene $\phi(c) = a$, $\phi(d) = b$. Haciendo $u = \phi(t)$ en la expresión anterior, se tiene $du = \phi'(t)\, dt = |\phi'(t)|\, dt$, de modo que

$$\longrightarrow = \int_a^b f(\lambda(u))\|\lambda'(u)\|\, du = \int_{\lambda} f\, ds$$

Si $\phi'(t) < 0\ \forall t \in [a, b]$, se tiene $\phi(c) = b$, $\phi(d) = a$. Haciendo $u = \phi(t)$ en la expresión (*), se tiene $du = \phi'(t)\, dt = -|\phi'(t)|\, dt$, de modo que

$$\longrightarrow = \int_b^a f(\lambda(u))\|\lambda'(u)\|(-du) = \int_a^b f(\lambda(u))\|\lambda'(u)\|\, du = \int_{\lambda} f\, ds \qquad \blacksquare$$

Insistimos en que en el teorema anterior se establece la propiedad de invariancia, *por reparametrizaciones del camino*, de la integral de línea respecto de la longitud de arco de la función f, en contraste con la propiedad análoga para integrales de línea de campos vectoriales establecida en el teorema 7.3.1, en la cual se establece la invariancia de las integrales de línea bajo composiciones de λ con funciones ϕ sobreyectivas (sin la propiedad de tener la derivada no nula, con la cual se producen reparametrizaciones de λ). De hecho, las integrales de línea con respecto a la longitud de arco tienen un carácter acumulativo en el sentido de que, si el recorrido de la traza del camino sufre algún "regreso", éste se acumula hacia el valor de la integral (como ocurrió con los ejemplos 5 y 6). Este tipo de recorridos con regresos múltiples no alteran el valor de la integral de línea de un campo vectorial, dependiendo dicho valor solamente del punto inicial, del punto final, y, claro está (en general), de la trayectoria que el camino sigue. Así pues, si λ es un camino de clase $\mathscr{C}^1$ que recorre su traza de manera inyectiva, y μ es un camino con la misma traza que λ, la cual recorre n veces (del punto inicial al final, de regreso al inicial, etc.) se tendrá

$$\int_{\mu} f\, ds = n \int_{\lambda} f\, ds$$

en tanto que para un campo vectorial $\mathbf{F}$ se tiene

$$\int_{\mu} \mathbf{F} \cdot d\mu = \begin{cases} 0 & \text{si } n \text{ es par} \\ \int_{\lambda} \mathbf{F} \cdot d\lambda & \text{si } n \text{ es impar} \end{cases}$$

En el ejemplo siguiente se muestra esta última afirmación.

Ejemplo 7. Sea $\lambda: [-1, 1] \to \mathbb{R}^n$ un camino de clase $\mathscr{C}^1$. Vamos a procurar un camino $\mu: [-1, 1] \to \mathbb{R}^n$ que tenga la misma traza que λ y que recorra ésta n veces. Para lograrlo, necesitamos componer λ con una función $\phi: [-1, 1] \to [-1, 1]$ con las siguientes características:

a. ϕ es de clase $\mathscr{C}^1$ y es sobreyectiva.

b. Se da una partición del intervalo $[-1, 1]$, digamos

$$-1 = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = 1$$

tal que

b1. la función ϕ manda cada subintervalo $[t_{i-1}, t_i]$ *sobre* $[-1, 1]$.

b2. $\phi'(t)$ es de signo constante en cada subintervalo (t_{i-1}, t_i).

Así, el camino $\boldsymbol{\mu} = \boldsymbol{\lambda} \circ \phi: [-1, 1] \to \mathbb{R}^n$ es de clase $\mathscr{C}^1$ y recorre la traza de $\boldsymbol{\lambda}$ n-veces, a saber, una vez en cada subintervalo $J_i = [t_{i-1}, t_i]$ de "ida" si $\phi'(t) > 0$ en (t_{i-1}, t_i), o de "regreso" si $\phi'(t) < 0$ en (t_{i-1}, t_i).

(**NOTA:** para los objetivos que persigue el presente ejemplo no es necesario hacer explícita la función ϕ, sino solamente usar sus propiedades. Sin embargo, aclaramos que, de hecho, hay funciones polinomiales con las características requeridas para ϕ. Por ejemplo, se puede demostrar que las funciones $T_n: [-1, 1] \to \mathbb{R}$ definidas como

$$T_n(x) = \cos(n \arccos x)$$

son funciones polinomiales de grado n, llamadas "polinomios de Chebychev", cuyo rango es el intervalo $[-1, 1]$, cuyos extremos absolutos en

$$x_k = \cos\left(\frac{k\pi}{n}\right), \quad k = 0, 1, \ldots, n$$

valen $f(x_k) = (-1)^k$, y por lo tanto, son máximos absolutos para k par y mínimos absolutos para k impar. Algunos de estos polinomios son $T_0(x) = 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, $T_4(x) = 8x^4 - 8x^2 + 1$, etc.

En los ejercicios al final de esta sección se hace una presentación detallada de los polinomios de Chebychev.)

Si $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función continua definida en el abierto U de $\mathbb{R}^n$ el cual contiene la imagen del camino $\boldsymbol{\lambda}$, se tiene

$$\begin{aligned}
\int_{\boldsymbol{\mu}} f\, ds &= \int_{-1}^{1} f(\boldsymbol{\mu}(t))\|\boldsymbol{\mu}'(t)\|\, dt = \int_{-1}^{1} f(\boldsymbol{\lambda}(\phi(t)))\|(\boldsymbol{\lambda} \circ \phi)'(t)\|\, dt \\
&= \int_{-1}^{1} f(\boldsymbol{\lambda}(\phi(t)))\|\boldsymbol{\lambda}'(\phi(t))\||\phi'(t)|\, dt \\
&= \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} f(\boldsymbol{\lambda}(\phi(t)))\|\boldsymbol{\lambda}'(\phi(t))\||\phi'(t)|\, dt
\end{aligned}$$

Si $\phi'(t) > 0$ en $[t_{i-1}, t_i]$, entonces $\phi(t_{i-1}) = -1$, $\phi(t_i) = 1$ de modo que

$$\int_{t_{i-1}}^{t_i} f(\boldsymbol{\lambda}(\phi(t)))\|\boldsymbol{\lambda}'(\phi(t))\||\phi'(t)|\, dt \overset{u=\phi(t)}{=\!=\!\longrightarrow}$$

$$= \int_{-1}^{1} f(\boldsymbol{\lambda}(u))\|\boldsymbol{\lambda}'(u)\|\, du = \int_{\boldsymbol{\lambda}} f\, ds$$

Si $\phi'(t) < 0$ en $[t_{i-1}, t_i]$, entonces $\phi(t_{i-1}) = 1$, $\phi(t_i) = -1$ de modo que

$$\int_{t_{i-1}}^{t_i} f\big(\boldsymbol{\lambda}(\phi(t))\big)\|\boldsymbol{\lambda}'(\phi(t))\||\phi'(t)|\,dt \overset{u=\phi(t)}{=\!=\!=} \int_{1}^{-1} f(\boldsymbol{\lambda}(u))\|\boldsymbol{\lambda}'(u)\|(-du)$$

$$= \int_{-1}^{1} f\big(\boldsymbol{\lambda}(u)\big)\|\boldsymbol{\lambda}'(u)\|\,du = \int_{\boldsymbol{\lambda}} f\,ds$$

Así la expresión para $\int_{\boldsymbol{\mu}} f\,ds$ queda como

$$\int_{\boldsymbol{\mu}} f\,ds = \sum_{i=1}^{n} \int_{\boldsymbol{\lambda}} f\,ds = n\int_{\boldsymbol{\lambda}} f\,ds$$

como queríamos probar. Dejamos al lector que dé un argumento similar que muestre que la integral de línea de un campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ a lo largo de $\boldsymbol{\mu}$ es igual a cero (si n es par) o coincide con la integral de $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ (si n es impar). ■

7.7.2 Aplicaciones

Como ya vimos en el ejemplo 2, conociendo la densidad lineal de un alambre en el espacio, digamos que dada por la función $\rho = \rho(x, y, z)$ (en gr/cm), y el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ en cuya imagen "se encuentra el alambre", entonces su masa total se calcula como

$$M = \int_{\boldsymbol{\lambda}} \rho\,ds$$

Ejemplo 8. Calculemos la masa total de un alambre cuya forma es la de la imagen de la hélice $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t, t)$, si la densidad en cada punto es proporcional al cuadrado de la distancia del punto al origen, valiendo 1 gr/cm en el punto inicial $\boldsymbol{\lambda}(0) = (1, 0, 0)$. Se tiene entonces la función densidad $\rho(x, y, z) = k(x^2 + y^2 + z^2)$, donde k es la constante de proporcionalidad, que se calcula usando que $\rho(1, 0, 0) = 1$, obteniéndose que $k = 1$. Entonces $\rho(x, y, z) = x^2 + y^2 + z^2$, y la masa total del alambre es

$$M = \int_{\boldsymbol{\lambda}} \rho\,ds = \int_0^{2\pi} \rho(\boldsymbol{\lambda}(t))\|\boldsymbol{\lambda}'(t)\|\,dt$$

$$= \int_0^{2\pi} (\cos^2 t + \operatorname{sen}^2 t + t^2)\sqrt{(-\operatorname{sen} t)^2 + (\cos t)^2 + 1}\,dt$$

$$= \int_0^{2\pi} (1 + t^2)\sqrt{2}\,dt = 2\sqrt{2}\pi(1 + \frac{4}{3}\pi^2)\text{ gr}$$ ■

Si tenemos un alambre en el plano, imagen del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, con función densidad $\rho = \rho(x, y)$ gr/cm, ésta es análoga a la presentada en la sección 5 del capítulo 6, donde se consideraron figuras planas en $\mathbb{R}^2$ (con área $A > 0$), lo que nos lleva a los siguientes resultados del comportamiento mecánico del alambre:

(*) Los momentos estáticos del alambre en relación a los ejes x e y, M_x y M_y respectivamente, se calculan como

$$M_x = \int_{\boldsymbol{\lambda}} y\rho(x, y)\,ds, \qquad M_y = \int_{\boldsymbol{\lambda}} x\rho(x, y)\,ds$$

(*) El centro de masa del alambre (de masa total M) se encuentra en el punto $(\bar{x}, \bar{y})$, donde

$$\bar{x} = \frac{M_y}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} x\rho(x, y)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y)\, ds}, \qquad \bar{y} = \frac{M_x}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} y\rho(x, y)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y)\, ds}$$

(*) El momento de inercia I_L del alambre respecto de un eje L, cuya distancia al punto (x, y) (perteneciente al alambre) que resulta de $\delta(x, y)$, es

$$I_L = \int_{\boldsymbol{\lambda}} \delta^2(x, y)\rho(x, y)\, ds$$

Si el alambre está en $\mathbb{R}^3$ y es imagen del camino $\boldsymbol{\lambda}\colon [a, b] \to \mathbb{R}^3$, con función densidad $\rho = \rho(x, y, z)$, los resultados análogos a los anteriores son los siguientes:

(*) Momentos estáticos respecto de los planos coordenados xy, xz, yz

$$M_{xy} = \int_{\boldsymbol{\lambda}} z\rho(x, y, z)\, ds, \quad M_{xz} = \int_{\boldsymbol{\lambda}} y\rho(x, y, z)\, ds, \quad M_{yz} = \int_{\boldsymbol{\lambda}} x\rho(x, y, z)\, ds$$

(*) Centro de masa en $(\bar{x}, \bar{y}, \bar{z})$, donde

$$\bar{x} = \frac{M_{yz}}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} x\rho(x, y, z)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y, z)\, ds}, \qquad \bar{y} = \frac{M_{xz}}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} y\rho(x, y, z)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y, z)\, ds}.$$

$$\bar{z} = \frac{M_{xy}}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} z\rho(x, y, z)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y, z)\, ds}$$

(*) Momento de inercia I_L respecto del eje L, cuya distancia al punto (x, y, z) del alambre es $\delta(x, y, z)$

$$I_L = \int_{\boldsymbol{\lambda}} \delta^2(x, y, z)\rho(x, y, z)\, ds$$

Ejemplo 9. Calculemos el centro de masa de un alambre homogéneo en forma del semicírculo $y = \sqrt{r^2 - x^2}$. Un camino cuya imagen es este semicírculo es $\boldsymbol{\lambda}\colon [0, \pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (r\cos t, r\operatorname{sen} t)$. Por razones de simetría, la abscisa $\bar{x}$ del centro de masa es $\bar{x} = 0$. Obtengamos la ordenada $\bar{y}$. Se tiene

$$\bar{y} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} y\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} ds} = \frac{1}{\pi r}\int_0^{\pi} (r\operatorname{sen} t)(r)\, dt = \frac{2r}{\pi}$$

Así entonces, el centro de masa del alambre está en $(0, 2r/\pi)$. ∎

Ejemplo 10. Calculemos los momentos de inercia del alambre homogéneo de densidad ρ_0 gr/cm, imagen del camino $\boldsymbol{\lambda}: [-a, a] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, \cosh t)$, respecto de los ejes coordenados. Se tiene

$$\begin{aligned} I_x &= \int_{\boldsymbol{\lambda}} y^2 \rho_0\, ds = \rho_0 \int_{-a}^{a} \cosh^2 t \sqrt{1 + \operatorname{senh}^2 t}\, dt = \rho_0 \int_{-a}^{a} \cosh^3 t\, dt \\ &= \rho_0 \int_{-a}^{a} (1 + \operatorname{senh}^2 t) \cosh t\, dt = \rho_0 \left[\operatorname{senh} t + \frac{1}{3} \operatorname{senh}^3 t \right]_{-a}^{a} \\ &= 2\rho_0 \operatorname{senh} a (1 + \frac{1}{3} \operatorname{senh}^2 a) \\ I_y &= \int_{\boldsymbol{\lambda}} x^2 \rho_0\, ds = \rho_0 \int_{-a}^{a} t^2 \sqrt{1 + \operatorname{senh}^2 t}\, dt = \rho_0 \int_{-a}^{a} t^2 \cosh t\, dt \\ &= \rho_0 \left[(t^2 + 2) \operatorname{senh} t - 2t \cosh t \right]_{-a}^{a} = 2\rho_0 \left[(a^2 + 2) \operatorname{senh} a - 2a \cosh a \right] \end{aligned}$$

Se puede verificar que $I_y < I_x$ independientemente del valor de $a > 0$. ¿Podría el lector interpretar este hecho en términos del significado físico del momento de inercia? ■

Ejemplo 11. Calculemos el centro de masa del alambre del ejemplo 8, en forma de la hélice $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t, t)$, cuya función densidad es $\rho(x, y, z) = x^2 + y^2 + z^2$. La masa total del alambre ya ha sido calculada y es $M = 2\sqrt{2}\pi(1 + \frac{4}{3}\pi^2)$ gr. Calculemos los momentos estáticos del alambre respecto de los planos coordenados

$$\begin{aligned} M_{xy} &= \int_{\boldsymbol{\lambda}} z\rho(x, y, z)\, ds = \int_0^{2\pi} t(\cos^2 t + \operatorname{sen}^2 t + t^2)\sqrt{2}\, dt \\ &= \sqrt{2} \int_0^{2\pi} t(1 + t^2)\, dt = 2\sqrt{2}\pi^2(1 + 2\pi^2) \\ M_{xz} &= \int_{\boldsymbol{\lambda}} y\rho(x, y, z)\, ds = \int_0^{2\pi} \operatorname{sen} t(\cos^2 t + \operatorname{sen}^2 t + t^2)\sqrt{2}\, dt \\ &= \sqrt{2} \int_0^{2\pi} \operatorname{sen} t(1 + t^2)\, dt = -4\sqrt{2}\pi^2 \\ M_{yz} &= \int_{\boldsymbol{\lambda}} x\rho(x, y, z)\, ds = \int_0^{2\pi} \cos t(\cos^2 t + \operatorname{sen}^2 t + t^2)\sqrt{2}\, dt \\ &= \sqrt{2} \int_0^{2\pi} \cos t(1 + t^2)\, dt = 4\sqrt{2}\pi \end{aligned}$$

Entonces el centro de masa del alambre está en el punto $(\bar{x}, \bar{y}, \bar{z})$ en donde

$$\bar{x} = \frac{M_{yz}}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} x\rho(x, y, z)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y, z)\, ds} = \frac{4\sqrt{2}\pi}{2\sqrt{2}\pi(1 + \frac{4}{3}\pi^2)} = \frac{2}{1 + \frac{4}{3}\pi^2}$$

$$\bar{y} = \frac{M_{xz}}{M} = \frac{\displaystyle\int_{\boldsymbol{\lambda}} y\rho(x, y, z)\, ds}{\displaystyle\int_{\boldsymbol{\lambda}} \rho(x, y, z)\, ds} = \frac{-4\sqrt{2}\pi^2}{2\sqrt{2}\pi(1 + \frac{4}{3}\pi^2)} = \frac{-2\pi}{1 + \frac{4}{3}\pi^2}$$

$$\bar{z} = \frac{M_{xy}}{M} = \frac{\int_{\lambda} z\rho(x, y, z)\, ds}{\int_{\lambda} \rho(x, y, z)\, ds} = \frac{2\sqrt{2}\pi^2(1 + 2\pi^2)}{2\sqrt{2}\pi(1 + \frac{4}{3}\pi^2)} = \frac{\pi(1 + 2\pi^2)}{1 + \frac{4}{3}\pi^2}$$

que corresponde aproximadamente al punto (0.141248, −0.443744, 4.60145). ■

Con una función continua $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el abierto U de $\mathbb{R}^n$, y un camino de clase $\mathscr{C}^1$ $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$, $\boldsymbol{\lambda}([a, b]) \subset U$, definimos el *valor medio de la función f sobre el camino $\boldsymbol{\lambda}$*, denotado por $\bar{f}$, como

$$\bar{f}_{\lambda} = \frac{1}{L_{\lambda}} \int_{\lambda} f\, ds$$

donde L_{λ} es la longitud del camino $\boldsymbol{\lambda}$, es decir

$$L_{\lambda} = \int_{\lambda} ds = \int_a^b \|\boldsymbol{\lambda}'(t)\|\, dt$$

El valor de $\bar{f}$ es un tipo de promedio de los valores que toma la función a lo largo del camino $\boldsymbol{\lambda}$. Compare esta definición con su análoga establecida para funciones de dos y tres variables con integrales dobles (sección 5, capítulo 6, apartado 6.5.3) y con integrales triples (sección 8, capítulo 6, apartado 6.8.3). Obsérvese que se trata de *la misma idea*, la misma incluso que se presenta en el primer curso de cálculo en el Teorema del Valor Medio para funciones reales de una variable real.

Ejemplo 12. El alambre del ejemplo 8, imagen de la hélice $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^3$ dada por $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t, t)$, tiene una función densidad $\rho(x, y, z) = x^2 + y^2 + z^2$. Calculemos la *densidad media* de este alambre, es decir, el valor medio de la función densidad a lo largo de $\boldsymbol{\lambda}$. La longitud de $\boldsymbol{\lambda}$ (digamos que en cm) es

$$L_{\lambda} = L_{\lambda} = \int_{\lambda} ds = \int_0^{2\pi} \|(-\operatorname{sen} t, \cos t, 1)\|\, dt = \int_0^{2\pi} \sqrt{2} = 2\sqrt{2}\pi$$

Por otra parte, la integral de la función densidad respecto de la longitud de arco a lo largo de $\boldsymbol{\lambda}$ es justamente la masa total del alambre que fue calculada en el ejemplo 8. Es decir

$$\int_{\lambda} \rho\, ds = 2\sqrt{2}\pi(1 + \frac{4}{3}\pi^2) \text{ gr}$$

Entonces la densidad media $\bar{\rho}$ del alambre es

$$\bar{\rho} = \frac{1}{L_{\lambda}} \int_{\lambda} \rho\, ds = 1 + \frac{4}{3}\pi^2 \text{ gr/cm}$$ ■

Ejemplo 13. Consideremos la función $f(x, y) = \sqrt{R^2 - x^2 - y^2}$ definida en $U = \{(x, y) | x^2 + y^2 \leq R^2\} \subset \mathbb{R}^2$. Sea $\boldsymbol{\lambda}: [-R, R] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, 0)$. El valor medio de f sobre $\boldsymbol{\lambda}$ es

$$\begin{aligned}\bar{f}_{\lambda} &= \frac{1}{L_{\lambda}} \int_{\lambda} f\, ds = \frac{1}{2R} \int_{-R}^{R} f(\boldsymbol{\lambda}(t)) \|\boldsymbol{\lambda}'(t)\|\, dt = \frac{1}{2R} \int_{-R}^{R} \sqrt{R^2 - t^2}\, dt \\ &= \frac{1}{2R} \text{ (área de un semicírculo de radio } R) = \frac{1}{2R}\left(\frac{\pi R^2}{2}\right) = \frac{\pi}{4} R\end{aligned}$$

(compare con el ejemplo 14 de la sección 5 del capítulo 6). ■

Otra interesante aplicación de las integrales de línea con respecto a la longitud de arco, la última que aquí discutiremos, está relacionada con el siguiente problema: supongamos una barda de altura constante h levantada sobre el plano xy, de modo que su base (la proyección de la barda sobre el plano xy), es una curva, imagen de un camino λ de clase $\mathscr{C}^1$ dado. Es claro que en este caso el área de la barda es el producto de la longitud del camino λ por la altura constante h. Supongamos ahora que la altura de la barda es variable, digamos que es una función f que en cada punto $\lambda(t) = (x, y)$ de la imagen del camino λ, asocia un valor real no negativo $f(\lambda(t)) =$ altura de la barda en $\lambda(t)$. En este caso, el área de la barda se puede calcular integrando la función altura f *sobre el camino* λ (a lo largo de él). Es decir, el área A es

$$A = \int_\lambda f\, ds$$

Figura 2. La integral de línea de f a lo largo de s, como un área.

Ejemplo 14. Imaginemos un teatro en forma de tetraedro, como el formado por los planos coordenados y el plano $z = \frac{b}{a}(a - x - y)$, donde $b > a > 0$, de modo que su base es un triángulo isóceles que tiene vértices en el origen y en los puntos $(a, 0, 0)$ y $(0, a, 0)$, y su altura es b. El escenario es circular y ocupa la región del plano xy dada por $\{(x, y)|x^2 + y^2 \leq r^2, x \geq 0,\ y\ \geq 0\}$ en donde $r < \sqrt{3}a/2$, como se muestra en la figura 3.

Sobre el límite del escenario se va a levantar una cortina hasta el techo del teatro. Se quiere saber qué área se cubrirá con la cortina. Aplicando matemáticas a este problema, vemos que se trata de integrar la función $f(x, y) = \frac{b}{a}(a-x-y)$ sobre el camino λ: $[0, \pi/2] \to \mathbb{R}^2$. $\lambda(t) = (r\cos t, r\operatorname{sen} t)$. Tenemos entonces que el área A procurada es

$$\begin{aligned} A &= \int_\lambda f\, ds = \int_0^{\pi/2} \frac{b}{a}(a - r\cos t - r\operatorname{sen} t)r\, dt \\ &= \frac{br}{a}\Big[at - r\operatorname{sen} t + r\cos t\Big]_0^{\pi/2} = \frac{br}{2a}(\pi a - 4r) \text{ unidades}^2 \end{aligned}$$ ■

Ejemplo 15. (El jardín de la catenaria). El Palacio de Mathingham es el lugar favorito de los matemáticos de todo el mundo para pasar unos días de descanso, conversando con sus colegas sobre

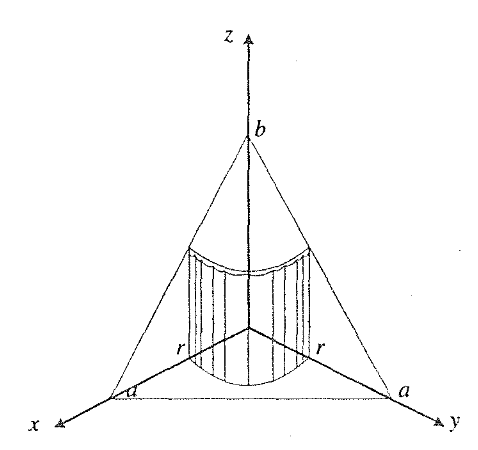


Figura 3. Figura del ejemplo 13.

los problemas no triviales que les ocupan en ese momento (es decir, a los que aún no les encuentran solución; los que ya tienen resueltos pasan automáticamente a la categoría de "triviales"). Para festejar el advenimiento del siglo XXI, el comité del Palacio ha decidido la construcción de un nuevo jardín para solaz de sus visitantes. El jardín deberá ser un homenaje a la bellísima catenaria, la curva que, se sabía desde los gloriosos tiempos de Johann Bernoulli, en el siglo del iluminismo, es la forma que toma una cadena por su propio peso al colgarse de dos puntos a la misma altura. Para dar su lugar al número 3 del milenio que comenzará con el siglo XXI, la forma que tendrá el remate superior de la fachada del jardín deberá ser una catenaria de la forma $v = 3\cosh\frac{u}{3}$. Así pues, el acceso al jardín estará rematado por arriba con la catenaria C, cuyo eje de simetría E será el eje de la puerta de entrada (de dos hojas). El mínimo local de C se encontrará a 3 m del piso. La puerta medirá un metro de cada lado y el frente en el que ésta se encuentra mide 4 m de cada lado de E. Habrá bardas laterales con altura igual a los máximos absolutos de C, cuyo fondo es de 15 m. La proyección sobre el piso de la barda final del jardín será una catenaria C' idéntica a C, con eje de simetría E', sobre la cual estará levantada la barda, cuya altura en el punto $\mathbf{p}'$ que dista d de E' deberá ser la misma que la de la barda en el punto $\mathbf{p}$ del frente del jardín, que dista d del eje E. Se trata entonces de calcular el área total de la barda que se habrá de levantar sobre la frontera del jardín. Coloquemos un sistema coordenado cartesiano como el mostrado en la figura 4.

El área A_1 de la barda frontal no es más que el área bajo la curva C (la cual está en el plano zy) que tiene por ecuación $z = 3\cosh\frac{y}{3}$ en la región $[-4, -1] \cup [1, 4]$ (es decir, de -4 a 4 eliminando el intervalo $[-1, 1]$ que ocupa la puerta). Aprovechando la simetría de la curva, calculamos A_1 como

$$\begin{aligned} A_1 &= 2\int_1^4 3\cosh\frac{y}{3}\,dy = 18\left[\operatorname{senh}\frac{y}{3}\right]_1^4 \\ &= 18\left(\operatorname{senh}\frac{4}{3} - \operatorname{senh}\frac{1}{3}\right) \approx 25.6589 \text{ m}^2 \end{aligned}$$

La altura de las bardas laterales del jardín es el valor de $z = 3\cosh\frac{y}{3}$ en $y = 4$, es decir $h = 3\cosh\frac{4}{3} \approx 6.08589$ m de modo que las áreas A_2 y A_3 de las bardas laterales del jardín son

$$A_2 = A_3 = \left(3\cosh\frac{4}{3}\right)(15) \approx 91.2884 \text{ m}^2$$

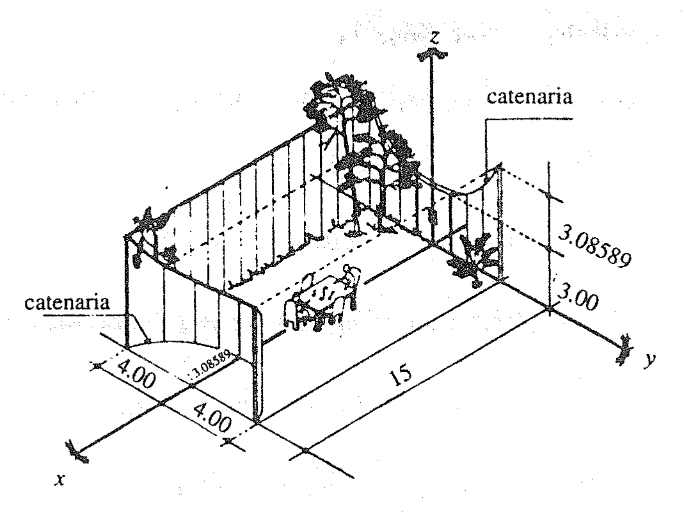


Figura 4. El jardín de la catenaria.

El punto interesante en este problema es el cálculo del área A_4 de la barda del fondo del jardín. Obsérvese que se trata de calcular la integral de línea respecto de la longitud de arco de la función $f(x, y) = 3\cosh\frac{y}{3}$ = altura de la barda en el punto (x, y), a lo largo del camino $\boldsymbol{\lambda}$ cuya imagen es la catenaria C'. Hallemos entonces una expresión para el camino $\boldsymbol{\lambda}$. La ecuación de C', que se encuentra en el plano xy, es la de una catenaria que pasa por los puntos $y = \pm 4$, $x = 15$, y cuyo "vértice" se encuentra a una distancia igual a $15 - 3.08589 = 11.9141$ del origen. Siendo la ecuación de la forma $x = \alpha\cosh\beta\, y$, tenemos que

$$x = 11.9141,\ y = 0 \Longrightarrow 11.9141 = \alpha\cosh 0 \Longrightarrow \alpha = 11.9141$$
$$x = 15,\ y = \pm 4 \Longrightarrow 15 = 11.9141\cosh 4\beta \Longrightarrow \beta \approx 0.176261$$

Es decir, la catenaria C' tiene por ecuación (en el plano xy) a

$$x = 11.9141\cosh(0.176261y)$$

con y variando de -4 a 4. Entonces el camino $\boldsymbol{\lambda}$ debe ser $\boldsymbol{\lambda}\colon[-4, 4] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (11.9141\cosh(0.176261t), t)$. El área A_4 de la barda final del jardín es entonces

$$\begin{aligned} A_4 &= \int_{\boldsymbol{\lambda}} f\, ds = \int_{-4}^{4} f(\boldsymbol{\lambda}(t))\|\boldsymbol{\lambda}'(t)\|\, dt \\ &= \int_{-4}^{4}\left(3\cosh\frac{t}{3}\right)\sqrt{1 + 4.41\,\mathrm{senh}^2(0.176261t)}\, dt \end{aligned}$$

El valor de esta integral (calculado numéricamente) equivaldrá a 43.7777. Así pues, el área de la barda que se debe levantar en la periferia del jardín de la catenaria debe ser (aproximadamente igual a)

$$A = 25.6589 + 2(91.2884) + 43.7777 = 252.0134\ \mathrm{m}^2$$ ■

Ejercicios (Capítulo 7, Sección 7)

En los ejercicios 1–10, calcule las integrales de línea con respecto a la longitud de arco indicadas.

1. $\int_\lambda x\,ds$, donde $\lambda: [0, 2] \to \mathbb{R}^2$, $\lambda(t) = (t, t)$.

2. $\int_\lambda (x + 3y)\,ds$, donde $\lambda: [-1, 1] \to \mathbb{R}^2$, $\lambda(t) = (t - 1, 3t + 2)$.

3. $\int_\lambda (x^2 - 2xy)\,ds$, donde $\lambda: [0, \pi/2] \to \mathbb{R}^2$, $\lambda(t) = (-\operatorname{sen} t, \cos t)$.

4. $\int_\lambda xy\,ds$, donde $\lambda: [0, \pi] \to \mathbb{R}^2$, $\lambda(t) = (3\cos t, 4\operatorname{sen} t)$.

5. $\int_\lambda (x^2 + y^2)\,ds$, donde $\lambda: [0, 2\pi] \to \mathbb{R}^2$, $\lambda(t) = (\operatorname{sen} t, \cos t)$.

6. $\int_\lambda (x + y)\,ds$, donde $\lambda: [-1, 1] \to \mathbb{R}^2$, $\lambda(t) = (t - 1, t^2)$.

7. $\int_\lambda y^2\,ds$, donde $\lambda: [0, 2] \to \mathbb{R}^2$, $\lambda(t) = (t, \cosh t)$.

8. $\int_\lambda (x - y + 3z)\,ds$, donde $\lambda: [0, 2] \to \mathbb{R}^3$, $\lambda(t) = (t + 1, t, 4t)$.

9. $\int_\lambda \frac{z}{x^2 + y^2}\,ds$, donde $\lambda: [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (\cos t, \operatorname{sen} t, t)$.

10. $\int_\lambda (3z + x^2 + y^2)\,ds$, donde $\lambda: [0, \pi] \to \mathbb{R}^3$, $\lambda(t) = (t\cos t, t\operatorname{sen} t, t)$.

11. Suponga que la integral de línea con respecto a la longitud de arco de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, a lo largo del camino $\lambda: [0, 1] \to \mathbb{R}^n$, $\lambda([0, 1]) \subset U$, es igual a k. Determine:

    a. $\int_\mu f\,ds$, donde $\mu: [0, 1] \to \mathbb{R}^n$, $\mu(t) = \lambda(1 - t)$.

    b. $\int_\mu f\,ds$, donde $\mu: [0, 1/2] \to \mathbb{R}^n$, $\mu(t) = \lambda(2t)$.

    c. $\int_\mu f\,ds$, donde $\mu: [0, 1] \to \mathbb{R}^n$, $\mu(t) = \lambda(t^2)$.

    d. $\int_\mu f\,ds$, donde $\mu: [0, 1] \to \mathbb{R}^n$, $\mu(t) = \lambda(\sqrt{t})$.

    e. $\int_\mu f\,ds$, donde $\mu: [0, \pi] \to \mathbb{R}^n$, $\mu(t) = \lambda(\operatorname{sen} t)$.

    f. $\int_\mu f\,ds$, donde $\mu: [0, 2\pi] \to \mathbb{R}^n$, $\mu(t) = \lambda(|\operatorname{sen} t|)$.

    g. $\int_\mu f\,ds$, donde $\mu: [-1, 1] \to \mathbb{R}^n$, $\mu(t) = \lambda(1 - t^2)$.

12. Suponga que $\int_{\boldsymbol{\lambda}} f\,ds = k_1$, $\int_{\boldsymbol{\lambda}} g\,ds = k_2$, donde $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^n$ es un camino cuya imagen está contenida en el conjunto abierto $U \subseteq \mathbb{R}^n$, dominio de las funciones f y g. Determine:

a. $\displaystyle\int_{\boldsymbol{\mu}} (f+g)\,ds$, en donde $\boldsymbol{\mu}: [0, 1] \to \mathbb{R}^n$, $\boldsymbol{\mu}(t) = \boldsymbol{\lambda}(1-t)$.

b. $\displaystyle\int_{\boldsymbol{\mu}} (3f-g)\,ds$, en donde $\boldsymbol{\mu}: [0, 2\pi] \to \mathbb{R}^n$, $\boldsymbol{\mu}(t) = \boldsymbol{\lambda}(|\cos t|)$.

c. $\displaystyle\int_{\boldsymbol{\mu}} (f-3g)\,ds$, en donde $\boldsymbol{\mu}: [-2\pi, 2\pi] \to \mathbb{R}^n$, $\boldsymbol{\mu}(t) = \boldsymbol{\lambda}(|\operatorname{sen} t|)$.

13. Halle la masa total de un alambre cuya forma es la de la curva $y = \ln x$, comprendida entre $x_1 = 1$ y $x_2 = e$, si la densidad en cada punto $\mathbf{p}$ de él es igual al cuadrado de la abscisa del punto.

14. Halle la masa total de un alambre "en v", cuya forma es la de la curva $y = |x|$, comprendida entre $x_1 = -1$ y $x_2 = 1$, si la densidad en cada punto de él es igual al valor absoluto del producto de las coordenadas del punto.

En los ejercicios 15–18, halle las coordenadas del centro de masa del alambre homogéneo cuya forma es la de la curva dada.

15. El cuadrado $|x| + |y| = 1$.

16. La mitad del cuadrado del ejercicio anterior, correspondiente a $y \geq 0$.

17. La porción de la catenaria $y = \cosh x$, correspondiente a $-1 \leq x \leq 1$.

18. Los lados del triángulo isósceles, con vértices en $A = (0, 0)$, $B = (2, 0)$, $C = (1, h)$.

19. Calcule el valor medio de la función $f(x, y) = \sqrt{R^2 - x^2 - y^2}$ sobre el camino $\boldsymbol{\lambda}: [-\alpha, \alpha] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, mt)$, en donde $\alpha = \frac{R}{\sqrt{1+m^2}}$ y m es un número real dado. (Sugerencia: en realidad no tiene que hacer ningún cálculo —pues ya están hechos—. Interprete geométricamente el problema y use el ejemplo 13 de esta sección).

20. Suponga que la curva, imagen del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^n$ de clase $\mathscr{C}^1$ es una curva de nivel de la función continua $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. ¿Cuál es el valor medio de la función f sobre $\boldsymbol{\lambda}$?

21. Calcular el área del cilindro $x^2 + y^2 = r^2$ que se encuentra dentro de la esfera $x^2 + y^2 + z^2 = R^2$, $(r < R)$.

22. Calcular el área del cilindro $x^2 + y^2 = ax$, que se encuentra dentro de la esfera $x^2 + y^2 + z^2 = a^2$.

23. Calcular el área del cilindro $x^2 + y^2 = ax$ que se encuentra por encima del plano $z = 0$ y por debajo de: **a.** $z = \sqrt{x^2 + y^2}$, **b.** $z = x^2 + y^2$.

24. Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función continua y positiva. Sea a un número del rango de f. Suponga que la curva del nivel a de f es una curva cerrada simple C, cuya longitud es L. ¿Cuál es el área del cilindro recto que se encuentra por debajo de la superficie $z = f(x, y)$ y por encima del plano $z = 0$, al cual corta en la curva C?

25. Sea $g: [a, b] \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ y sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función continua y positiva, en donde $g([a, b]) \subset U$. Demuestre que el área de la parte del cilindro $y = g(x)$ que se encuentra por encima del plano $z = 0$ y por debajo de la superficie $z = f(x, y)$ es igual al producto de la longitud de la gráfica de $y = g(x)$, $a \leq x \leq b$, por el valor medio de la función f sobre el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, g(t))$.

26. (Un ejercicio sobre polinomios de Chebychev). Para cada $n = 0, 1, 2, \ldots$ defina la función $T_n: [-1, 1] \to \mathbb{R}$ como

$$T_n(x) = \cos(n \arccos x)$$

A este tipo de funciones se les conoce como "polinomios de Chebychev" (más en concreto, $T_n(x)$ es el polinomio de Chebychev de grado n) y tienen muchos usos en problemas de aproximación e interpolación.

a. Haga $y = \arccos x$, y escriba $T_n(\cos y) = \cos(ny)$. Con ayuda de las identidades trigonométricas del coseno de la suma y diferencia de dos ángulos concluya que

$$T_{n+1}(\cos y) + T_{n-1}(\cos y) = 2\cos(ny)\cos y$$

Obtenga entonces que

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

para $n \geq 1$.

b. Con el resultado del inciso anterior demuestre que $T_n(x)$ es un polinomio de grado n. (A pesar de que este polinomio está definido sólo en $[-1, 1]$, veremos a continuación que "las cosas importantes" de él, por ejemplo sus raíces y sus extremos, ocurren todas dentro de su dominio). Compruebe que el coeficiente de x^n en $T_n(x)$ es 2^{n-1}.

c. Compruebe que $T_0(x) = 1$, $T_1(x) = X$. Usando la fórmula obtenida en el inciso a, obtenga los polinomios de Chebychev $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, $T_4(x) = 8x^4 - 8x^2 + 1$, $T_5(x) = 16x^5 - 20x^3 + 5x$. Bosqueje las gráficas de estos polinomios.

d. Demuestre que si n es par, $T_n(x)$ es una función par, mientras que si n es impar, $T_n(x)$ será impar.

e. Demuestre que para $n \geq 1$, el polinomio de Chebychev $T_n(x)$ tiene n ceros simples (de multiplicidad uno) en

$$x_k = \cos\left(\frac{2k-1}{2n}\pi\right), \quad k = 1, 2, \ldots, n$$

(todas dentro del dominio $[-1, 1]$ de $T_n(x)$).

f. Demuestre que para $n \geq 2$, el polinomio de Chebychev $T_n(x)$ tiene $n - 1$ puntos críticos en

$$x_k^* = \cos\left(\frac{k\pi}{n}\right), \quad k = 1, 2, \ldots, n-1$$

(todos dentro del dominio $[-1, 1]$ de $T_n(x)$).

g. Use el resultado del inciso anterior para probar que $T_n(x)$ tiene extremos absolutos en

$$x_k^* = \cos\left(\frac{k\pi}{n}\right), \quad k = 0, 1, 2, \ldots, n$$

los cuales son máximos absolutos iguales a 1 si k es par, y mínimos absolutos iguales a -1 si k es impar.

7.8 La perspectiva de la física

En la sección 4 se han planteado y resuelto algunos cuestionamientos interesantes sobre las integrales de línea. Sin embargo, una parte muy importante de la riqueza de esta teoría se descubre hasta que se estudian los aspectos físicos que se describen con esta herramienta matemática. El objetivo de esta sección es presentar algunos de estos aspectos físicos que requieren de la matemática desarrollada en este capítulo para su estudio, como el concepto de trabajo y el estudio de los campos gravitacionales y eléctricos. Veremos que la nomenclatura empleada con las funciones de la sección 4 ("funciones potenciales", "campos conservativos") queda justificada cuando se usan en el contexto físico los resultados ahí estudiados.

A manera de introducción, recordemos cómo se presenta el concepto de trabajo en los cursos de física elemental: la popular fórmula que servía para definir este concepto era "trabajo igual a fuerza por distancia", lo cual significaba que al actuar una fuerza f sobre un cuerpo, éste se desplazaba una distancia d (en la dirección sobre la que actuaba f); el trabajo W realizado sobre el cuerpo lo daba $W = fd$.

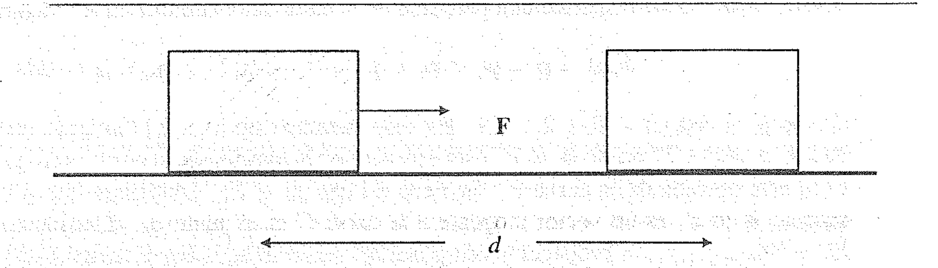


Figura 1. El concepto de trabajo en la física elemental.

Consideremos ahora un problema mucho más complicado. Supongamos que en cada punto $\mathbf{p} = (x, y)$ de una región U de $\mathbb{R}^2$ actúa una fuerza $\mathbf{F}$ cuyo valor depende del punto $\mathbf{p}$. Escribimos entonces $\mathbf{F} = \mathbf{F}(x, y)$. Tenemos así un campo vectorial en el conjunto U de $\mathbb{R}^2$, $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, *el cual es un campo de fuerzas.* Si tomamos un objeto "puntual", *en el punto* $\mathbf{p}$, y *queremos llevarlo, a través del campo* $\mathbf{F}$, a *lo largo de una trayectoria dada, del punto* $\mathbf{p}$ a *otro punto* $\mathbf{q}$ *de* U, debemos, obviamente, realizar un "trabajo", análogo al de "fuerza por distancia" al mover el cuerpo en el ejemplo preliminar. El asunto es que ahora la situación es terriblemente más complicada que la de este ejemplo elemental. Ahora el trabajo no será más "fuerza por distancia", porque, en principio, la fuerza $\mathbf{F}$ está cambiando punto a punto de la trayectoria que describe el cuerpo.

Para aplicar matemáticas a la solución de este problema, vamos a suponer que nuestro campo de fuerzas es lo suficientemente "bien portado" desde el punto de vista diferenciable como para que las operaciones que hagamos a continuación tengan sentido. Supondremos además que la trayectoria sobre la que se moverá el cuerpo a través del campo está dada por (la traza de) un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, cuya imagen (que denotaremos por C) está contenida en U, el conjunto en que se define el campo $\mathbf{F}$. Este camino también se supone tan diferenciable en cuanto se necesite. Además, se tiene que $\boldsymbol{\lambda}(a) = \mathbf{p}$, $\boldsymbol{\lambda}(b) = \mathbf{q}$.

Tomemos una partición del intervalo $[a, b]$ como sigue

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b$$

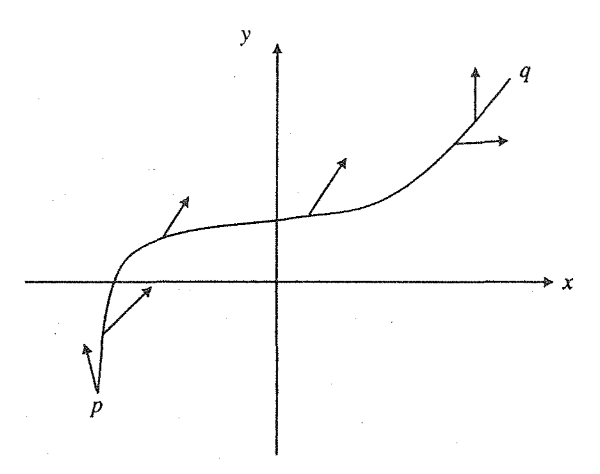


Figura 2. Un punto moviéndose a través de un campo **F**, del punto **p** al punto **q**.

A esta partición corresponde una partición de la traza del camino $\boldsymbol{\lambda}$ en $\mathbb{R}^2$, digamos

$$\boldsymbol{\lambda}(a) = \mathbf{p} = \mathbf{p}_0 < \mathbf{p}_1 < \mathbf{p}_2 < \cdots < \mathbf{p}_{n-1} < \mathbf{p}_n = \mathbf{q} = \boldsymbol{\lambda}(b)$$

donde $\mathbf{p}_i = \boldsymbol{\lambda}(t_i)$, $i = 0, 1, 2, \ldots, n$. En cada subintervalo $[t_{i-1}, t_i]$ tomemos un punto arbitrario ξ_i. Sea $\boldsymbol{\rho}_i = \boldsymbol{\lambda}(\xi_i)$. El punto $\boldsymbol{\rho}_i \in \mathbb{R}^2$ está entonces en la porción de la curva entre $\mathbf{p}_{i-1}$ y $\mathbf{p}_i$. Llamemos C_i a esta porción de la curva C. Se tiene así que $\boldsymbol{\rho}_i \in C_i$. Obsérvese que $\boldsymbol{\lambda}'(\xi_i)$, la derivada del camino $\boldsymbol{\lambda}$ en ξ_i, es un vector tangente a la curva C en el punto $\boldsymbol{\rho}_i$. Consideremos ahora al vector $\tilde{\mathbf{F}}_i = \mathrm{PR}_{\mathbf{F}(\boldsymbol{\rho}_i)\to\boldsymbol{\lambda}'(\xi_i)}$ = proyección ortogonal del vector $\mathbf{F}(\boldsymbol{\rho}_i)$ sobre el vector $\boldsymbol{\lambda}'(\xi_i)$. Descomponiendo el vector $\mathbf{F}(\boldsymbol{\rho}_i)$ como la suma de $\tilde{\mathbf{F}}_i$ y su otra componente ortogonal, resulta claro que *solamente el vector* $\tilde{\mathbf{F}}_i$ *es la "parte" del vector* $\mathbf{F}(\boldsymbol{\rho}_i)$ *responsable del trabajo que se efectúa en el recorrido del punto* $\mathbf{p}$ *sobre la curva* C, pues ésta es la parte *en la dirección tangencial* (que marca la dirección del desplazamiento del punto que recorre la curva) del vector fuerza $\mathbf{F}(\boldsymbol{\rho}_i)$.

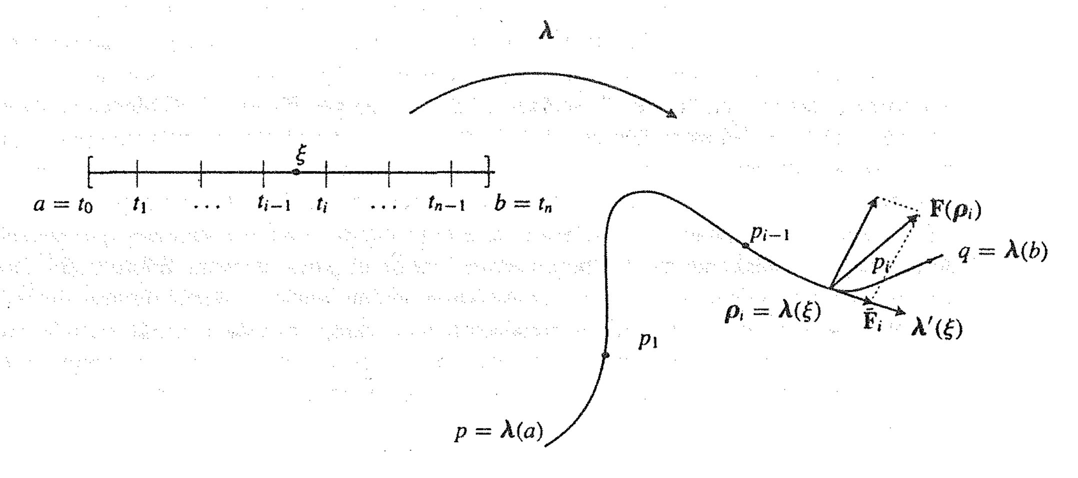


Figura 3. Una partícula de la curva C y la descomposición del campo **F** en la suma de sus componentes ortogonales, una de las cuales está en la dirección tangencial a la curva.

Más aún, en la porción C_i de la curva C, podemos hacer una aproximación del trabajo realizado al recorrer ese tramo de curva, multiplicando la magnitud de la fuerza $\tilde{\mathbf{F}}_i$ por la distancia lineal $\|\mathbf{p}_i - \mathbf{p}_{i-1}\|$. Se tiene entonces que el trabajo procurado, que denotaremos por $W_{\mathbf{pq}}$, es aproximadamente

$$W_{\mathbf{pq}} \approx \sum_{i=1}^{n} \|\tilde{\mathbf{F}}_i\| \, \|\mathbf{p}_i - \mathbf{p}_{i-1}\|$$

Esta aproximación será tanto mejor cuanto n sea grande y los puntos $\mathbf{p}_i$ estén más cerca uno de otro. Recuerde que

$$\tilde{F}_i = \mathrm{PR}_{F(\rho_i)\to\boldsymbol{\lambda}'(\xi_i)} = \frac{F(\rho_i)\cdot\boldsymbol{\lambda}'(\xi_i)}{\|\boldsymbol{\lambda}'(\xi)_i\|^2}\boldsymbol{\lambda}'(\xi_i)$$

(ver secciones 2 y 3 del capítulo 1). La magnitud de esta fuerza es entonces

$$\|\tilde{\mathbf{F}}_i\| = \frac{\mathbf{F}(\rho_i)\cdot\boldsymbol{\lambda}'(\xi_i)}{\|\boldsymbol{\lambda}'(\xi)_i\|}$$

(donde el numerador es el producto punto del vector $\mathbf{F}(\rho_i)$ por el vector $\boldsymbol{\lambda}'(\xi_i)$). Por otra parte, la distancia lineal entre el punto $\mathbf{p}_i = \boldsymbol{\lambda}(t_i)$ y el punto $\mathbf{p}_{i-1} = \boldsymbol{\lambda}(t_{i-1})$ es aproximadamente igual a la longitud del camino $\boldsymbol{\lambda}$ en el intervalo $[t_{i-1}, t_i]$. Es decir que

$$\|\mathbf{p}_i - \mathbf{p}_{i-1}\| \approx \int_{t_i}^{t_{i-1}} \|\boldsymbol{\lambda}'(t)\|\,dt \approx \|\boldsymbol{\lambda}'(\xi_i)\|(t_i - t_{i-1})$$

donde el último signo $\approx$ se sigue del Teorema del Valor Medio para integrales aplicado a la función $\|\boldsymbol{\lambda}'(t)\|$ en el intervalo $[t_{i-1}, t_i]$[1]

Juntando todas las piezas, llegamos a que una aproximación del valor del trabajo $W_{\mathbf{pq}}$ es

$$\begin{aligned} W_{pq} &\approx \sum_{i=1}^{n} \|\tilde{F}_i\| \|\mathbf{p}_i - \mathbf{p}_{i-1}\| \\ &\approx \sum_{i=1}^{n} \frac{F(\rho_i)\cdot\boldsymbol{\lambda}'(\xi_i)}{\|\boldsymbol{\lambda}'(\xi_i)\|}\|\boldsymbol{\lambda}'(\xi_i)\|(t_i - t_{i-1}) \\ &= \sum_{i=1}^{n} (F(\rho_i)\cdot\boldsymbol{\lambda}'(\xi_i))(t_i - t_{i-1}) \end{aligned}$$

Al tomar el límite cuando n tiende a infinito y la norma de la partición original tiende a cero, obtenemos el valor del trabajo $W_{\mathbf{pq}}$ procurado. Obsérvese que la última expresión se identifica

---

[1] Según el teorema del valor medio, se tiene que

$$\int_{t_i}^{t_{i-1}} \|\boldsymbol{\lambda}'(t)\|\,dt = \|\boldsymbol{\lambda}'(\zeta)\|(t_i - t_{i-1})$$

donde $t_{i-1} < \zeta < t_i$. Si consideramos que el punto ζ del que habla este teorema es el punto ξ_i que ya habíamos tomado en el intervalo $[t_{i-1}, t_i]$, ya no se tiene la igualdad en la fórmula que establece el teorema. Es claro, sin embargo, que si t_{i-1} está tan cerca de t_i, que los puntos ζ y ξ_i no pueden estar muy lejos uno del otro, tiene entonces un resultado aproximado en la fórmula mencionada.

fácilmente como la suma de Riemann de la función ϕ: $[a, b] \to \mathbb{R}$, $\phi(t) = F(t) \cdot \boldsymbol{\lambda}'(t)$. Así se tiene que

$$W_{pq} = \lim_{\substack{n \to \infty \\ \text{norma de } P \to 0}} \sum_{i=1}^{n} (F(\rho_i) \cdot \boldsymbol{\lambda}'(\xi_i))(t_i - t_{i-1})$$
$$= \int_a^b F(t) \cdot \boldsymbol{\lambda}'(t)\, dt$$

la cual es ni más ni menos que la integral de línea del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}$. Esta discusión, hecha para un campo en $\mathbb{R}^2$, es igualmente válida para un campo en $\mathbb{R}^3$ y sirve para campos continuos y caminos seccionalmente $\mathscr{C}^1$. En resumen se tiene lo siguiente:

Sea $\mathbf{F}$: $U \subseteq \mathbb{R}^2$ (o $\mathbb{R}^3$) $\to \mathbb{R}^2$ (o $\mathbb{R}^3$) un campo de clase $\mathscr{C}^k$, $k \geq 0$ y sea $\boldsymbol{\lambda}$: $[a, b] \to \mathbb{R}^2$ (o $\mathbb{R}^3$) un camino seccionalmente $\mathscr{C}^1$ cuya imagen está contenida en U. El trabajo que hay que realizar para llevar un cuerpo del punto $\mathbf{p} = \boldsymbol{\lambda}(a)$ al punto $\mathbf{q} = \boldsymbol{\lambda}(b)$ por el camino $\boldsymbol{\lambda}$ a través del campo $\mathbf{F}$ es igual al valor de la integral de línea del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}$. Es decir

$$W_{\mathbf{pq}} = \int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda}$$

Si reconsideramos la situación elemental planteada inicialmente, con una fuerza constante f actuando sobre un cuerpo, el cual es desplazado una distancia d, tendríamos el campo $\mathbf{F}$: $\mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (f, 0)$ y el camino $\boldsymbol{\lambda}$: $[0, d] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, 0)$, de modo que el trabajo para mover el cuerpo una distancia d es

$$W = \int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} = \int_0^d ((f)(t)' + 0(0)')\, dt = f \int_0^d dt = fd$$

que es la fórmula de la física elemental que comentábamos al principio de esta sección.

Ejemplo 1. Consideremos el campo de fuerzas $\mathbf{F}$: $\mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y, y)$. El trabajo que hay que realizar para llevar un punto desde el origen de coordenadas $\mathbf{p} = (0, 0)$ hasta el punto $\mathbf{q} = (1, 1)$ a través de este campo, por la curva $y = x^2$, es decir, a lo largo del camino $\boldsymbol{\lambda}$: $[0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$, es

$$W_{\mathbf{pq}} = \int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} = \frac{4}{3}$$

(ver ejemplo 1 de la sección 3). ■

Usemos el ejemplo anterior para reflexionar sobre algunas propiedades de las integrales de línea estudiadas en la sección 3. Es claro que si tomamos una reparametrización $\tilde{\boldsymbol{\lambda}}$ del camino $\boldsymbol{\lambda}$, que conserve su orientación (es decir, que $\tilde{\boldsymbol{\lambda}}$ recorra en la misma dirección la curva $y = x^2$, con, posiblemente, otra velocidad), el valor del trabajo calculado en el ejemplo anterior no debe cambiar, pues éste finalmente depende sólo del campo de fuerzas involucrado y la curva que conecta el punto inicial y el punto final de la trayectoria. Este es el resultado establecido en el teorema 7.2.1, **a**, sobre la invariancia de las integrales de línea por reparametrizaciones del camino. Más aún, si invertimos el sentido del recorrido, yendo ahora del punto $\mathbf{q}$ al punto $\mathbf{p}$, es de esperar que esto se refleje en el valor del trabajo de alguna manera, ya que no es lo mismo un recorrido sobre una trayectoria "en el sentido del campo —con las flechas apuntando a favor del recorrido— que en dirección contraria". Esta idea puede resultar más clara si vemos el esquema en donde el que muestran las "flechas" del campo que actúan en los puntos de la trayectoria $y = x^2$ (ver figura 4). De hecho, sabemos que

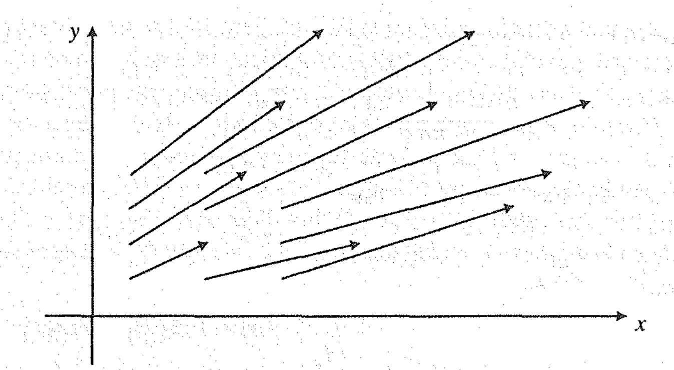


Figura 4. El campo del ejemplo 1.

si invertimos el sentido de recorrido, esto se reflejará en un cambio de signo en la integral de línea (teorema 7.2.1).

Adentrémonos más en el contenido físico de las integrales de línea: consideremos un campo de fuerzas $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ y el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ cuya imagen está en U. Tomemos una partícula de masa m que viajará por el camino $\boldsymbol{\lambda}$ de $\mathbf{p} = \boldsymbol{\lambda}(a)$ a $\mathbf{q} = \boldsymbol{\lambda}(b)$ a través del campo $\mathbf{F}$. El trabajo realizado para efectuar tal viaje es, como hemos visto,

$$W_{\mathbf{pq}} = \int_{\boldsymbol{\lambda}} F \cdot d\boldsymbol{\lambda} = \int_a^b F(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\, dt$$

Según la 2a. ley de Newton, el valor de la fuerza $\mathbf{F}$ en el punto $\boldsymbol{\lambda}(t)$ es igual a la masa m de la partícula por su aceleración en ese punto. Esta última es, como se vio en el capítulo de curvas, el vector segunda derivada $\boldsymbol{\lambda}''(t)$. Entonces

$$W_{\mathbf{pq}} = \int_a^b F(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\, dt = m \int_a^b \boldsymbol{\lambda}''(t) \cdot \boldsymbol{\lambda}'(t)\, dt$$

Recuerde que (ver capítulo 5, sección 3)

$$\frac{d}{dt}\|\boldsymbol{\lambda}'(t)\|^2 = 2\boldsymbol{\lambda}''(t) \cdot \boldsymbol{\lambda}'(t)$$

de modo que

$$W_{\mathbf{pq}} = \frac{m}{2} \int_a^b \frac{d}{dt}\|\boldsymbol{\lambda}'(t)\|^2\, dt = \frac{1}{2}m\|\boldsymbol{\lambda}'(b)\|^2 - \frac{1}{2}m\|\boldsymbol{\lambda}'(a)\|^2$$

Esta es una bonita fórmula que repetimos a continuación

$$W_{\mathbf{pq}} = \frac{1}{2}m\|\boldsymbol{\lambda}'(b)\|^2 - \frac{1}{2}m\|\boldsymbol{\lambda}'(a)\|^2$$

Observe que $\|\boldsymbol{\lambda}'(t)\|$ es la magnitud del vector velocidad de la partícula en el instante t. Escribamos $v(\mathbf{p}) = \|\boldsymbol{\lambda}'(a)\|$, $v(\mathbf{q}) = \|\boldsymbol{\lambda}'(b)\|$, para denotar tales magnitudes del vector velocidad (correspondientes a los puntos $\mathbf{p}$ y $\mathbf{q}$, en los instantes $t = a$ y $t = b$, respectivamente). La fórmula anterior se ve entonces como

$$\boxed{W_{\mathbf{pq}} = \frac{1}{2}m(v(\mathbf{q}))^2 - \frac{1}{2}m(v(\mathbf{p}))^2}$$

Los sumandos involucrados en la fórmula anterior (un medio del producto de la masa de la partícula por el cuadrado de su velocidad) se identifican fácilmente como la *energía cinética* de la partícula de masa m. Hemos establecido entonces que *el trabajo necesario para mover una partícula de masa m por el camino* $\boldsymbol{\lambda}$ *del punto* $\mathbf{p} = \boldsymbol{\lambda}(a)$ *al punto* $\mathbf{q} = \boldsymbol{\lambda}(b)$ a *través del campo* $\mathbf{F}$ *es igual a la diferencia entre la energía cinética de la partícula en el punto* $\mathbf{q}$ y *la correspondiente en el punto* $\mathbf{p}$.

Supongamos ahora que el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ es un campo conservativo. Sea $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ una función potencial de $\mathbf{F}$, es decir, tal que grad $f(x, y, z) = F(x, y, z)$, para $(x, y, z) \in U$. Como se vio en la demostración del teorema 7.4.1, el valor de la integral de línea de $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ es

$$\int_{\lambda} F \cdot d\boldsymbol{\lambda} = f(\boldsymbol{\lambda}(b)) - f(\boldsymbol{\lambda}(a))$$

Comparando esta expresión con la obtenida anteriormente para $W_{\mathbf{pq}}$ en términos de la energía cinética de la partícula se obtiene que

$$\frac{1}{2}m(v(\mathbf{q}))^2 - \frac{1}{2}m(v(\mathbf{p}))^2 = f(\boldsymbol{\lambda}(b)) - f(\boldsymbol{\lambda}(a))$$

o bien (escribiendo $\mathbf{p} = \boldsymbol{\lambda}(a)$, $\mathbf{q} = \boldsymbol{\lambda}(b)$)

$$\boxed{\frac{1}{2}m(v(\mathbf{p}))^2 - f(\mathbf{p}) = \frac{1}{2}m(v(\mathbf{q}))^2 - f(\mathbf{q})}$$

Al escalar $-f(\xi)$ se le llama en física *energía potencial* (de la partícula en el punto $\xi \in \mathbb{R}^3$). De esta manera se justifica que a la función $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ tal que grad $f = F$, se le llame *función potencial*. La fórmula anterior establece entonces que la suma de la energía cinética y la energía potencial de la partícula *es la misma* (es decir *se conserva*) en los puntos inicial y final de la trayectoria. De hecho, esta suma es la misma en *cualquier* punto intermedio de la trayectoria de la partícula (¿por qué?). Este es el conocido *principio de conservación de la energía*, y por esta razón a los campos para los que es válido tal principio se les llama campos *conservativos*.

Consideremos ahora, a manera de ejemplo, el campo gravitacional producido por un cuerpo de masa M (ver sección 2). Coloquemos un objeto de masa M en el origen de coordenadas en $\mathbb{R}^3$. En cada punto $(x, y, z) \in \mathbb{R}^3 - \{(0, 0, 0)\}$, un objeto de masa unitaria será atraído por M, atendiendo a la Ley de Gravitación Universal de Newton, con una fuerza cuya magnitud es directamente proporcional al producto de la masa M y la masa $m = 1$, e inversamente proporcional al cuadrado de la distancia que los separa. Puesto que M está en el origen, esta distancia es $\|(x, y, z)\|$. Tenemos así un campo vectorial (el campo gravitacional producido por la masa M) dado por $\mathbf{F}: \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}^3$, que

$$\|F(x, y, z)\| = \frac{GM}{\|(x, y, z)\|^2}$$

donde G es la constante de gravitación universal. Entonces el campo $\mathbf{F}$ debe ser

$$F(x, y, z) = -\frac{GM}{\|(x, y, z)\|^3}(x, y, z)$$

(el signo menos indica que el vector $\mathbf{F}(x, y, z)$ va del punto (x, y, z) al origen). Veamos que este campo es conservativo. Para esto, exhibiremos una función potencial $f: \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}$, de

modo que grad $f = F$. Tal función es

$$f(x, y, z) = \frac{GM}{\|(x, y, z)\|}$$

Se tiene, en efecto,

$$\frac{\partial f}{\partial x}(x, y, z) = \frac{\partial}{\partial x}\left(\frac{GM}{\sqrt{x^2 + y^2 + z^2}}\right) = -\frac{GM}{(x^2 + y^2 + z^2)^{3/2}}x$$

obteniéndose expresiones similares para las derivadas parciales de f con respecto a y y a z. Entonces

$$\begin{aligned}\operatorname{grad} f(x, y, z) &= \left(\frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z)\right)\\ &= -\frac{GM}{(x^2 + y^2 + z^2)^{3/2}}(x, y, z) = -\frac{GM}{\|(x, y, z)\|^3}(x, y, z) = F(x, y, z)\end{aligned}$$

Con ayuda de la función potencial f podemos establecer una fórmula para el trabajo de llevar la masa unitaria m de un punto $r_1 = (x_1, y_1, z_1)$ a otro punto $r_2 = (x_2, y_2, z_2)$ a través del campo gravitacional producido por M. Puesto que el campo es conservativo, el valor de tal trabajo no depende de la trayectoria que conecta r_1 con r_2. Según vimos anteriormente, este valor es igual a la diferencia de la función potencial en r_2 con el correspondiente en r_1 (teorema 7.4.1). Así entonces

$$W_{r_1 r_2} = f(r_2) - f(r_1) = \frac{GM}{\|r_2\|} - \frac{GM}{\|r_1\|}$$

Consideraremos ahora otro ejemplo muy importante de campo vectorial análogo al campo gravitacional discutido anteriormente: el campo eléctrico producido por una carga q, digamos negativa. Colocando esta carga en el origen de $\mathbb{R}^3$, (la magnitud de) el campo eléctrico producido por ella en el punto $\mathbf{p} = (x, y, z) \in \mathbb{R}^3 - \{(0, 0, 0)\}$ se obtiene como la (magnitud de la) fuerza F_p dividida por q_0, en la que q_0 es una carga de prueba positiva colocada en el punto $\mathbf{p}$, la cual, según la ley de Coulomb, experimentará una fuerza F_p (en este caso, de atracción) hacia la carga q que se encuentra en el origen, siendo (la magnitud de) esta fuerza directamente proporcional al producto de la carga q por la carga q_0 e inversamente proporcional al cuadrado de la distancia que las separa (es decir, $\|(x, y, z)\|$). Entonces el campo eléctrico producido por la carga q es $E\colon \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}^3$

$$E(x, y, z) = -\frac{kq}{\|(x, y, z)\|^3}(x, y, z)$$

donde k es la constante de proporcionalidad de la ley de Coulomb. Nótese de nuevo, por analogía con el campo gravitacional analizado anteriormente, que este campo es conservativo, pues la función $\phi\colon \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}$, dada por

$$\phi(x, y, z) = \frac{kq}{\|(x, y, z)\|}$$

es una función potencial para el campo E, y que el trabajo que hay que realizar para llevar la carga q_0 de un punto $r_1 = (x_1, y_1, z_1)$ a otro punto $r_2 = (x_2, y_2, z_2)$ a través de este campo es

$$W_{r_1 r_2} = \phi(r_2) - \phi(r_1) = \frac{kq}{\|r_2\|} - \frac{kq}{\|r_1\|}$$

Al valor de la función potencial ϕ en el punto $r = (x, y, z)$ se le llama *potencial* del campo eléctrico en ese punto, el cual se denota como V_r. En electrostática se define la *diferencia de potencial* entre los puntos A y B de un campo eléctrico, denotada por V_{AB}, como el trabajo que hay que realizar para llevar una carga unitaria positiva, a través del campo, *del punto A al punto B* (ver, por ejemplo, el libro [*Ht*], pp. 87, 88, 89). Escribiendo $A = r_1$ y $B = r_2$, y usando la fórmula establecida anteriormente para el trabajo $W_{r_1 r_2}$, tenemos entonces que

$$V_{AB} = W_{r_2 r_1} = \phi(r_1) - \phi(r_2) = \frac{kq}{\|r_1\|} - \frac{kq}{\|r_2\|} = V_A - V_B$$

Así pues, en este caso (campo eléctrico producido por una carga q situada en el origen de coordenadas), la diferencia de potencial V_{AB} resulta ser efectivamente una *diferencia* entre los potenciales de los puntos A y B.

Ejercicios (Capítulo 7, Sección 8)

1. Calcular el trabajo que se necesita para llevar un punto material de $(1, 0)$ a $(-1, 0)$, por el círculo $x^2 + y^2 = 1$, a través de un campo de fuerzas en el que en cada punto del plano actúa una fuerza constante de magnitud igual a 2, apuntando en la dirección positiva del eje y.
2. Repetir el problema anterior si el punto se lleva por la curva $y = 1 - |x|$.
3. En cada punto del plano actúa una fuerza de magnitud constante igual a 2, la cual forma siempre proyecciones iguales con los ejes coordenados. Calcular el trabajo necesario para desplazar un punto material a través de este campo de fuerzas, del punto $\mathbf{p} = (0, 0)$, al punto $\mathbf{q} = (2, 2)$, en cada uno de los siguientes casos:
   a. yendo por el segmento de recta que une a $\mathbf{p}$ con $\mathbf{q}$.
   b. siguiendo la trayectoria (de segmentos de recta): $\mathbf{p} = (0, 0) \to (1, 0) \to (1, 1) \to (2, 1) \to (2, 2) = \mathbf{q}$.
   c. yendo por la parábola $y = x^2/2$.
4. Considere el campo de fuerzas $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x + y, x - y)$. Calcule el trabajo que se necesita para trasladar un punto material a través de este campo desde el origen de coordenadas hasta el punto $(2, 0)$ en cada uno de los casos siguientes:
   a. yendo directamente sobre el eje x.
   b. siguiendo la trayectoria del semicírculo $(x - 1)^2 + y^2 = 1$, $y \geq 0$.
   c. siguiendo la trayectoria del semicírculo $(x - 1)^2 + y^2 = 1$, $y \leq 0$.
5. Considere el campo de fuerzas $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x, y, z)$. Calcular el trabajo que se necesita para llevar un punto a través de este campo, desde el origen de coordenadas hasta el punto $(1, 1, 1)$, en cada uno de los siguientes casos:
   a. yendo directamente por un segmento de recta.
   b. siguiendo la (de segmentos de recta) trayectoria $(0, 0, 0) \to (1, 0, 0) \to (1, 1, 0) \to (1, 1, 1)$.
6. Un cuerpo de masa $m = 1$ gr se mueve a través de un campo de fuerzas siguiendo el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$. Suponga que $\boldsymbol{\lambda}'(a) = (1, 2, 1)$, $\boldsymbol{\lambda}'(b) = (1, 3, 2)$. Calcule el trabajo que se realiza para llevar el cuerpo de $\mathbf{p} = \boldsymbol{\lambda}(a)$ a $\mathbf{q} = \boldsymbol{\lambda}(b)$.

7. Verifique la ley de la conservación de la energía con un cuerpo de masa $m = 1$ gr moviéndose a través del campo de fuerzas $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (y, x)$, por el camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^3)$. Es decir, una vez comprobado que el campo $\mathbf{F}$ es conservativo, compruebe que la suma de la energía cinética más la energía potencial en el punto $\mathbf{p} = \boldsymbol{\lambda}(a)$ es igual a la suma de la energía cinética más la energía potencial en el punto $\mathbf{q} = \boldsymbol{\lambda}(b)$.

8. Repita el ejercicio anterior con un cuerpo de masa $m = 1$ gr, el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (3y^2z + ye^x, 6xyz + e^x, 3xy^2)$ y el camino $\boldsymbol{\lambda}: [0, 1] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (t, t, t)$. (Ver ejemplo 13 de la sección 4).

7.9 Teorema de Green

En esta sección estudiaremos uno de los resultados clásicos del cálculo en $\mathbb{R}^n$ el cual relaciona integrales de línea con integrales dobles. Trabajaremos con campos en $\mathbb{R}^2$, $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ de clase $\mathscr{C}^1$, digamos $\mathbf{F}(x, y) = (M(x, y), N(x, y))$ y con cierto tipo de regiones $S \subset U$, que llamaremos "compactas". La forma más general de este resultado, así como su demostración, involucra una serie de detalles técnicos que escapan a los alcances del rigor de este libro. Sin embargo, quisiéramos dejar establecido el teorema en su forma más general y luego ver algunos casos particulares importantes de él, junto con sus demostraciones.

Dejando por el momento los detalles a un lado, el teorema de Green establece la siguiente igualdad entre una integral de línea y una integral doble

$$\int_{\partial S_+} F \cdot d\boldsymbol{\lambda} = \iint_S \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\, dy$$

donde $\mathbf{F} = (M, N)$, S es una cierta región de $\mathbb{R}^2$ y ∂S^+ es la frontera de S (imagen del camino $\boldsymbol{\lambda}$) positivamente orientada, lo cual significa que ésta se recorre "en el sentido contrario al de las manecillas del reloj" (esta idea no es muy precisa como veremos más adelante, pero es aceptable en una primera instancia), como se muestra en la figura 1.

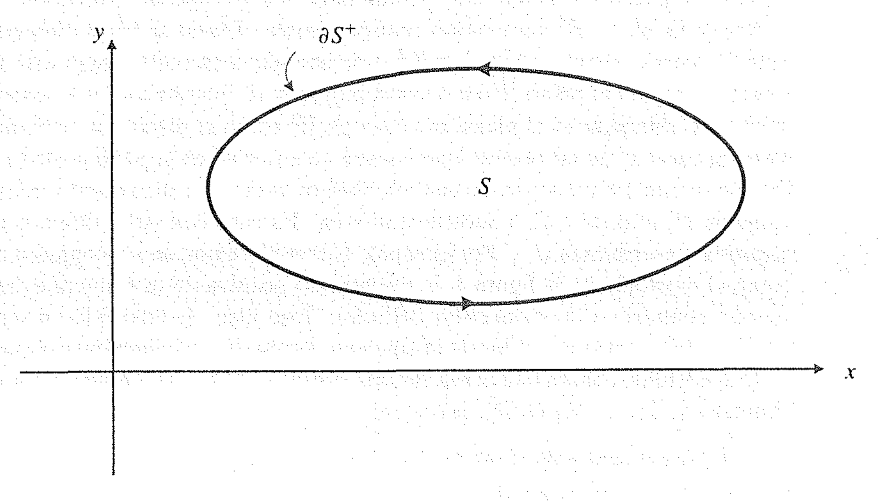


Figura 1. La región $S \subseteq \mathbb{R}^2$ y su frontera ∂S^+ positivamente orientada.

Las dificultades técnicas para precisar el enunciado de este teorema comienzan al tratar de definir con rigor el tipo de regiones $S \subseteq \mathbb{R}^2$ en las cuales el resultado es válido, así como lo que significa que su frontera esté positivamente orientada. Ciertamente la región S mostrada en la figura 1 es una de las regiones a las que se refiere el teorema, pero, por ejemplo, la región mostrada en la figura 2 cae también dentro de las regiones en las que el resultado es válido.

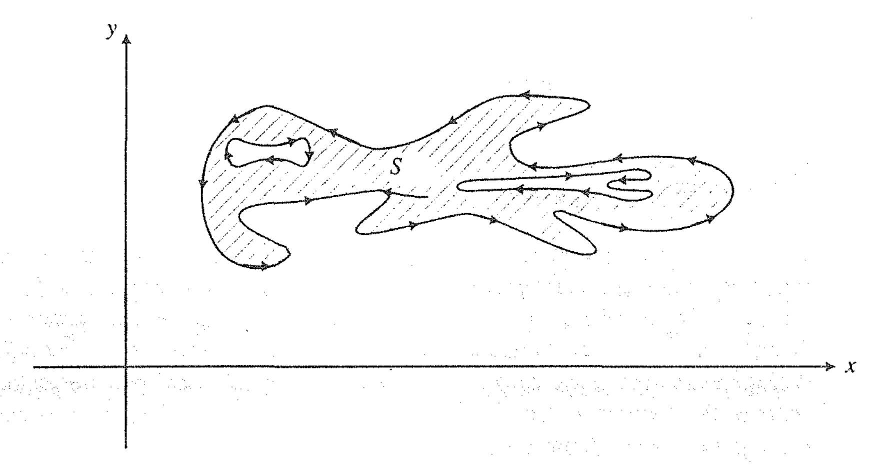


Figura 2. Una región S en la que el Teorema de Green es válido.

Así pues, es necesario establecer primeramente lo que significa "región compacta de $\mathbb{R}^2$", las cuales serán las regiones en las que es válido el Teorema de Green. En principio, debemos aceptar como verdadero el ya mencionado "Teorema de la Curva de Jordan" (ver la parte final de la sección 4), el cual establece que "toda curva cerrada simple C en el plano separa a éste en dos subconjuntos, uno acotado, llamado interior de C, y otro no acotado, llamado exterior de C, de los cuales la curva C es la frontera común". Como se había mencionado, este resultado, que parece obvio para algunos casos sencillos de curvas cerradas simples (que podemos imaginar de primera instancia: círculos, elipses, etc.), es un resultado muy profundo de la Topología de Conjuntos.

Sea $\lambda: [a, b] \to \mathbb{R}^2$ un camino cerrado simple. Como siempre, llamaremos a la traza de este camino "curva cerrada simple" en $\mathbb{R}^2$, y la denotaremos por C. Según el Teorema de la Curva de Jordan, C separa al plano $\mathbb{R}^2$ en los subconjuntos U, interior a C, y V, exterior a C. Imaginemos a la curva C dibujada en el plano xy, visto en $\mathbb{R}^3$ como el plano $z = 0$ e imáginemonos caminando sobre la curva C de tal manera que nuestra cabeza esté en la parte positiva del eje z. Diremos que C tiene *orientación positiva*, lo cual escribimos como C^+, si en nuestro recorrido siempre vemos al conjunto U, interior a C, a nuestra izquierda. En caso contrario, diremos que C tiene *orientación negativa* y escribimos C^-. Por ejemplo, si consideramos un subconjunto acotado convexo en $\mathbb{R}^2$, como el mostrado en la figura 1, la orientación positiva de su frontera es aquélla que la recorre en sentido contrario a las manecillas del reloj. Esta idea, de recorrido en sentido antihorario, ya no funciona con curvas como las de la siguiente figura, la cual también está positivamente orientada.

Consideremos ahora una curva cerrada simple C en $\mathbb{R}^2$, así como un conjunto de curvas cerradas simples $C_1, C_2, \ldots, C_k$ en $\mathbb{R}^2$, tales que:

a. C_i se encuentra en el interior de C, $i = 1, 2, \ldots, k$.

b. $C_i \cap C_j = \varnothing$, para $i \neq j$.

c. C_i se encuentra en el exterior de C_j, para $i \neq j$.

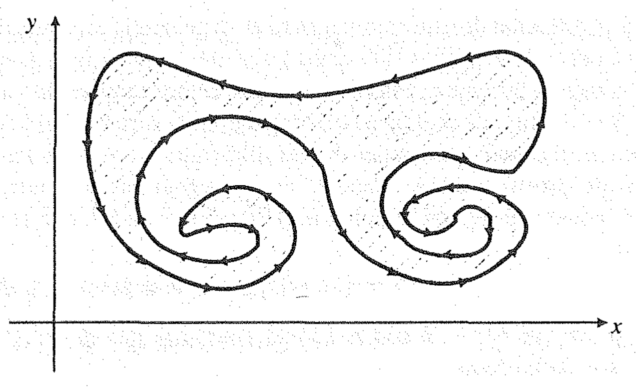


Figura 3. Una curva cerrada simple positivamente orientada.

Al subconjunto S formado por los puntos $(x, y) \in \mathbb{R}^2$ que se encuentran en el interior de C y en el exterior de cada una de las C_i, $i = 1, 2, \ldots, k$, junto con los mismos puntos (x, y) de las curvas C, $C_1, C_2, \ldots, C_k$, se le llama *región compacta* en $\mathbb{R}^2$. Un ejemplo de este tipo de región es el que se muestra en la figura 2. La *frontera* de la región compacta S, denotada por ∂S, es (por definición) el conjunto de las curvas $C, C_1, C_2, \ldots, C_k$ que la constituyen. Diremos que la frontera ∂S de la región compacta S está *positivamente orientada*, lo cual escribimos como ∂S^+, cuando al recorrer las curvas $C, C_1, C_2, \ldots, C_k$ de su frontera (según se convino anteriormente, con uno mismo caminando sobre las curvas con la cabeza en la parte positiva del eje z), la región S se mantiene siempre a la izquierda. De nuevo, en la figura 2 se muestra la región compacta S con su frontera positivamente orientada. Otro ejemplo se encuentra en la siguiente figura.

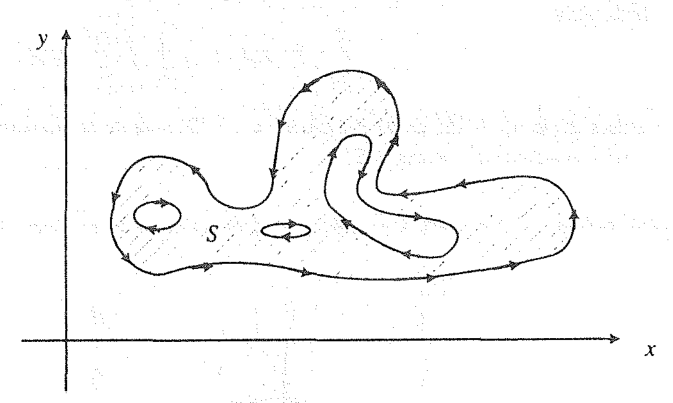


Figura 4. Una región compacta en $\mathbb{R}^2$ con su frontera positivamente orientada.

Ahora estamos en condiciones de enunciar el célebre teorema que da nombre a esta sección.

Teorema **Teorema de Green** Sea $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (M, N)$, un campo de clase $\mathscr{C}^1$ definido en el abierto U de $\mathbb{R}^2$. Sea $S \subset U$ una región compacta con su frontera ∂S^+ positivamente orientada. Entonces

$$\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_S\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dx\, dy$$

donde $\boldsymbol{\lambda}$ es un camino seccionalmente $\mathscr{C}^1$ cuya traza es ∂S^+. ∎

La demostración de este teorema en toda su generalidad, como decíamos anteriormente, cae fuera de los alcances de este libro. Sin embargo, veremos ahora la demostración del teorema para regiones S más o menos decentes. Más en concreto, demostraremos la validez de este resultado cuando la región S es del tipo de regiones consideradas en la sección 3 del capítulo 6 (con las que ya tenemos experiencia al calcular integrales dobles sobre ellas, como la del miembro derecho de la fórmula del teorema de Green), a saber, cuando S es una región que es al mismo tiempo del tipo I y del tipo II. Recordemos que éstas son regiones acotadas del plano $\mathbb{R}^2$, que se dicen ser del tipo I cuando son de la forma

$$S = \{(x, y)|a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$

donde $\phi_1, \phi_2: [a, b] \to \mathbb{R}$ son funciones continuas que $\phi_1(x) \leq \phi_2(x)$, $\forall x \in [a, b]$, y del tipo II cuando son de la forma

$$D_{\text{II}} = \{(x, y)|\psi_1(y) \leq x \leq \psi_2(y), c \leq y \leq d\}$$

donde $\psi_1, \psi_2: [c, d] \to \mathbb{R}$ dos funciones continuas, de modo que $\psi_1(y) \leq \psi_2(y)$, $\forall y \in [c, d]$. Comencemos por establecer un par de resultados previos a la demostración del teorema.

Teorema 7.9.1 Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función real definida en el abierto U de $\mathbb{R}^2$, de clase $\mathscr{C}^1$, y sea $S \subset U$ la región (del tipo I)

$$S = \{(x, y)|a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$

Entonces

$$\int_{\partial S^+} F \cdot d\boldsymbol{\lambda} = -\int_S\int \frac{\partial f}{\partial y}\, dx\, dy$$

donde $\mathbf{F}: U \subseteq\to \mathbb{R}^2$ es el campo $\mathbf{F} = (f, 0)$, y $\boldsymbol{\lambda}$ es un camino cuya traza es la frontera de S positivamente orientada, ∂S^+.

Demostración. La región S entonces es como la que se muestra en la figura 5.

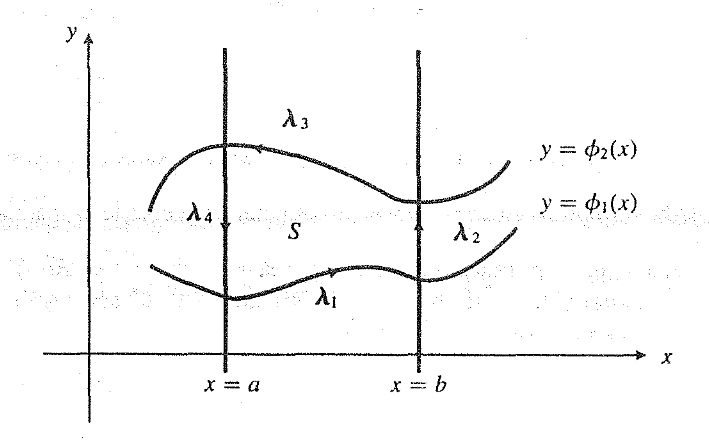


Figura 5. Región $S \subseteq \mathbb{R}^2$ del tipo I.

El camino $\boldsymbol{\lambda}$ se puede tomar como $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3 + \boldsymbol{\lambda}_4$, de modo que las trazas de cada uno de estos caminos $\boldsymbol{\lambda}_i$ sea uno de los cuatro "lados" de la frontera ∂S^+. Digamos que

$$\boldsymbol{\lambda}_1: [a, b] \to \mathbb{R}^2,\ \boldsymbol{\lambda}_1(t) = (t, \phi_1(t)).(\boldsymbol{\lambda}_1 \text{ recorre el lado inferior})$$

$$\boldsymbol{\lambda}_2: [\phi_1(b), \phi_2(b)] \to \mathbb{R}^2,\ \boldsymbol{\lambda}_2(t) = (b, t).(\boldsymbol{\lambda}_2 \text{ recorre el lado derecho})$$

$$\boldsymbol{\lambda}_3: [a, b] \to \mathbb{R}^2,\ \boldsymbol{\lambda}_3(t) = (a + b - t, \phi_2(a + b - t)).(\boldsymbol{\lambda}_3 \text{ recorre el lado superior})$$

$$\boldsymbol{\lambda}_4: [\phi_1(a), \phi_2(a)] \to \mathbb{R}^2,\ \boldsymbol{\lambda}_4(t) = (a, \phi_1(a) + \phi_2(a) - t).(\boldsymbol{\lambda}_4 \text{ recorre el lado izquierdo}).$$

Se tiene entonces

$$\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \sum_{i=1}^{4} \int_{\boldsymbol{\lambda}_i} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

(cada uno de los caminos $\boldsymbol{\lambda}_i$ es de clase $\mathscr{C}^1$; el camino $\boldsymbol{\lambda}$ es, por lo tanto, seccionalmente $\mathscr{C}^1$). Observe que si $i = 2$ o 4, se tiene

$$\int_{\boldsymbol{\lambda}_i} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\alpha}^{\beta} \big((f(\boldsymbol{\lambda}_i(t)))(0) + (0)(\pm 1)\big)\, dt = 0$$

(donde $\alpha = \phi_1(b)$, $\beta = \phi_2(b)$, si $i = 2$, o $\alpha = \phi_1(a)$, $\beta = \phi_2(a)$, si $i = 4$), de modo que

$$\int_{\partial S_+} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} \mathbf{F} \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_3} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

Se tiene

$$\begin{aligned}
\int_{\boldsymbol{\lambda}_1} \mathbf{F} \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_3} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_a^b f(t, \phi_1(t))\, dt + \int_a^b f(a + b - t, \phi_2(a + b - t))(-1)\, dt \\
&= \int_a^b f(t, \phi_1(t))\, dt - \int_a^b f(t, \phi_2(t))\, dt \\
&= -\int_a^b \Big(f(t, \phi_2(t)) - f(t, \phi_1(t))\Big)\, dt = -\int_a^b \left(\int_{\phi_1(t)}^{\phi_2(t)} \frac{\partial f}{\partial y}\, dy\right) dt \\
&= -\iint_S \frac{\partial f}{\partial y}\, dx\, dy
\end{aligned}$$

Así pues, se tiene

$$\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} = -\iint_S \frac{\partial f}{\partial y}\, dx\, dy$$

como se quería. ■

Teorema 7.9.2 Sea $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función real definida en el abierto U de $\mathbb{R}^2$, de clase $\mathscr{C}^1$, y sea $S \subset U$ la región (del tipo II)

$$S = \{(x, y) | \psi_1(y) \leq x \leq \psi_2(y), c \leq y \leq d\}$$

Entonces

$$\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} = \iint_S \frac{\partial f}{\partial x}\, dx dy$$

donde $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ es el campo $\mathbf{F} = (0, f)$, y $\boldsymbol{\lambda}$ es un camino cuya traza es la frontera de S positivamente orientada, ∂S^+.

Demostración. El argumento es análogo al del teorema anterior, así que obviamos los comentarios y escribimos sólo lo necesario. En este caso la región S se ve como

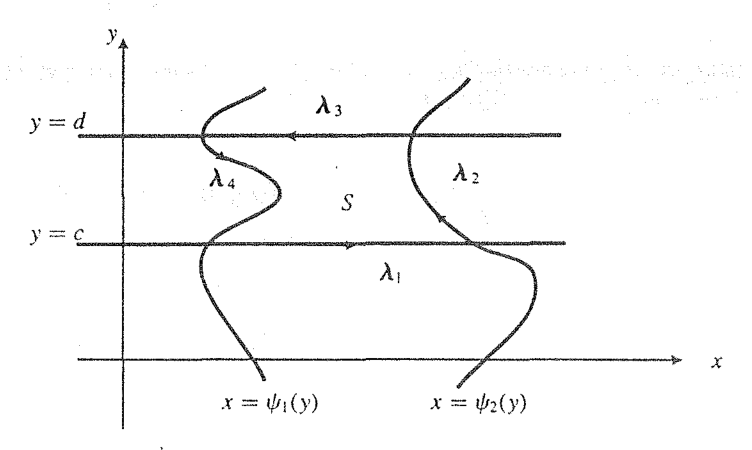


Figura 6. Región $S \subseteq \mathbb{R}^2$ del tipo II.

Se tiene $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3 + \boldsymbol{\lambda}_4$ donde

$$\boldsymbol{\lambda}_1: [\psi_1(c), \psi_2(c)] \to \mathbb{R}^2, \boldsymbol{\lambda}_1(t) = (t, c).$$
$$\boldsymbol{\lambda}_2: [c, d] \to \mathbb{R}^2, \boldsymbol{\lambda}_2(t) = (\psi_2(t), t).$$
$$\boldsymbol{\lambda}_3: [\psi_1(d), \psi_2(d)] \to \mathbb{R}^2, \boldsymbol{\lambda}_3(t) = (\psi_1(d) + \psi_2(d) - t, d).$$
$$\boldsymbol{\lambda}_4: [c, d] \to \mathbb{R}^2, \boldsymbol{\lambda}_4(t) = (\psi_1(c + d - t), c + d - t).$$

Entonces

$$\begin{aligned}
\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \sum_{i=1}^{4} \int_{\boldsymbol{\lambda}_i} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_2} \mathbf{F} \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_4} \mathbf{F} \cdot d\boldsymbol{\lambda} \\
&= \int_c^d f(\psi_2(t), t)\, dt + \int_c^d f(\psi_1(c + d - t), c + d - t)(-1)\, dt \\
&= \int_c^d f(\psi_2(t), t)\, dt - \int_c^d f(\psi_1(t), t)\, dt = \int_c^d \left(\int_{\psi_1(t)}^{\psi_2(t)} \frac{\partial f}{\partial x}\, dx \right) dt
\end{aligned}$$

$$= \iint_S \frac{\partial f}{\partial x}\, dxdy \qquad \text{Q.E.D.}$$

Ahora estamos en condiciones de dar la prueba del teorema de Green para regiones que son del tipo I y del tipo II a la vez.

Demostración. (del teorema de Green para regiones compactas S que son del tipo I y II). Escribamos el campo $\mathbf{F} = (M, N)$ como $\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2$, donde $\mathbf{F}_1 = (M, 0)$, $\mathbf{F}_2 = (0, N)$. Entonces, usando la linealidad de las integrales de línea y los resultados de los dos teoremas anteriores (de modo más preciso, el teorema 7.8.1 para el campo F_1 viendo S como región del tipo I, y el teorema 7.8.2 para el campo F_2 viendo a S como región del tipo II), tenemos

$$\int_{\partial S^+} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\partial S^+} (\mathbf{F}_1 + \mathbf{F}_2)\cdot d\boldsymbol{\lambda} = \int_{\partial S^+} \mathbf{F}_1\cdot d\boldsymbol{\lambda} + \int_{\partial S^+} \mathbf{F}_2\cdot d\boldsymbol{\lambda}$$
$$= -\iint_S \frac{\partial M}{\partial y}\, dxdy + \iint_S \frac{\partial N}{\partial x}\, dxdy = \iint_S \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dxdy \qquad \text{Q.E.D.}$$

Veamos algunos ejemplos en los que verificamos la validez del teorema de Green.

Ejemplo 1. Verifiquemos el teorema de Green con el campo $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ dado por $\mathbf{F}(x, y) = (3x^2y, -x^3)$ y S la región comprendida entre la parábola $y = x^2$ y la recta $y = 1$. En esquema se tiene

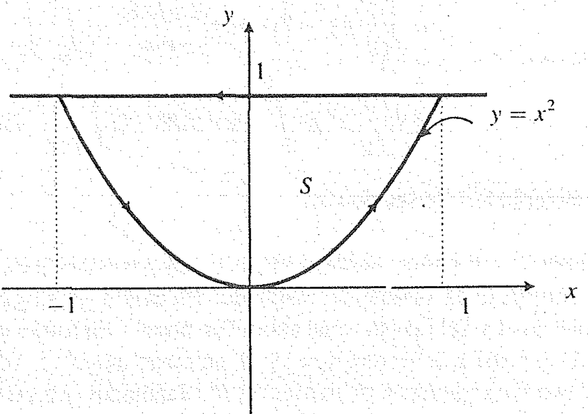


Figura 7. La región S del ejemplo 1.

La frontera de S orientada positivamente es la traza del camino $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$ donde $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2\colon [-1, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_1(t) = (t, t^2)$, $\boldsymbol{\lambda}_2(t) = (-t, 1)$. Entonces, por una parte se tiene

$$\int_{\partial S^+} \mathbf{F}\cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} \mathbf{F}\cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_2} \mathbf{F}\cdot d\boldsymbol{\lambda}$$
$$= \int_{-1}^{1} \big(3(t^2)(t^2)(1) + (-t^3)(2t)\big)\, dt + \int_{-1}^{1} \big(3(-t)^2(1)(-1)\big)\, dt$$
$$= \int_{-1}^{1} t^4\, dt - 3\int_{-1}^{1} t^2\, dt = \left[\frac{1}{5}t^5\right]_{-1}^{1} - \left[t^3\right]_{-1}^{1} = \frac{2}{5} - 2 = -\frac{8}{5}$$

Por otra parte

$$\int_S\int\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\int_{-1}^{1}\int_{x^2}^{1}\left(-3x^2-3x^2\right)dx\,dy$$
$$=\int_{-1}^{1}-6x^2(1-x^2)\,dx=\left[-2x^3+\frac{6}{5}x^5\right]_{-1}^{1}=-\frac{8}{5}$$

queda así establecida la validez de la fórmula del teorema de Green. ■

Ejemplo 2. Usemos el teorema de Green para calcular la integral de línea del campo $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (2x + y, -x + 4xy)$ sobre el círculo $x^2 + y^2 = 1$ recorrido (una vez) en sentido antihorario. Esta curva es la frontera orientada positivamente de la región

$$S = \{(x, y)|x^2 + y^2 \leq 1\}$$

Se tiene entonces que

$$\int_{\partial S^+}\mathbf{F}\cdot d\boldsymbol{\lambda}=\iint_S\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\iint_S\big((-1+4y)-(1)\big)\,dx\,dy$$
$$=\iint_S(4y-2)\,dx\,dy\overset{\text{coordenadas}}{\underset{\text{polares}}{\longrightarrow}}=\int_0^1\int_0^{2\pi}(4r\operatorname{sen}\theta-2)r\,dr\,d\theta$$
$$=4\int_0^1 r^2dr\int_0^{2\pi}\operatorname{sen}\theta\,d\theta-2\int_0^1\int_0^{2\pi}d\theta=4\int_0^1 r^2dr(0)-2\pi=-2\pi$$

(ver ejemplo 8 de la sección 3). ■

Podemos considerar regiones un poco más generales que éstas y convencernos, con un argumento muy simple, de la validez del resultado. En efecto, supongamos que la región $R \subseteq \mathbb{R}^2$, siendo a la vez del tipo I y del tipo II, tiene ahora "un hoyo". En forma más precisa, consideremos dos regiones R y R' que son a la vez del tipo I y II, de modo que $R' \subset R$. Sea S el conjunto de los puntos (x, y) de R que *no pertenecen* al interior de R' (denotados por Int R', definido como Int $R' = R' - \partial R'$). Es decir, $S = R - \operatorname{Int} R'$. Entonces S es una región compacta en $\mathbb{R}^2$ cuya frontera positivamente orientada es $\partial S^+ = \partial R^+ \cup \partial R'^-$ (ver figura 8). Sea $\boldsymbol{\lambda}$ un camino cuya imagen es ∂R^+, $\boldsymbol{\mu}$ un camino cuya imagen es $\partial R'^+$ y $\boldsymbol{\nu}$ un camino cuya imagen es $\partial R'^-$. Denotemos por ρ el camino cuya imagen es ∂S^+ (el cual es, en cierto sentido, una "unión" de los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\nu}$).

Si $\mathbf{F}\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (M, N)$, es un campo de clase $\mathscr{C}^1$, $U \supset R$, entonces

$$\iint_S\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dxdy=\iint_r\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dxdy-\iint_{R'}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dxdy$$
$$\overset{teoremadeGreen}{=\!=\!\longrightarrow}=\int_{\partial R^+}\mathbf{F}\cdot d\boldsymbol{\lambda}-\int_{\partial R'^+}\mathbf{F}\cdot d\boldsymbol{\mu}$$
$$=\int_{\partial R^+}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{\partial R'^-}\mathbf{F}\cdot d\boldsymbol{\nu}=\int_{\partial S^+}\mathbf{F}\cdot d\rho$$

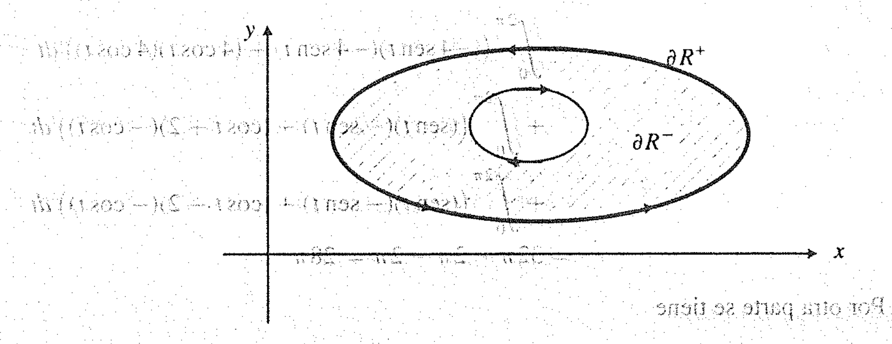


Figura 8. La región $S = R - \text{int } R'$ con su frontera positivamente orientada.

lo que muestra la validez del teorema de Green para la región compacta R. Este mismo argumento se puede usar para demostrar el teorema de Green en regiones compactas que son del tipo I y II "con un número finito de hoyos del tipo I y del tipo II".

Ejemplo 3. Consideremos las regiones

$$R_1 = \{(x, y)|x^2 + y^2 \leq 16\}$$
$$R_2 = \{(x, y)|(x - 2)^2 + y^2 \leq 1\}$$
$$R_3 = \{(x, y)|(x + 2)^2 + y^2 \leq 1\}$$

Sea $S = R_1 - (\text{Int } R_2 \cup \text{Int } R_3)$. S es entonces una región compacta de $\mathbb{R}^2$ y su frontera positivamente orientada ∂S^+ consiste en la unión de la frontera positivamente orientada de R_1 con las fronteras negativamente orientadas de R_2 y R_3. Es decir, $\partial S^+ = \partial R_1^+ \cup \partial R_2^- \cup \partial R_3^-$ (ver figura 9).

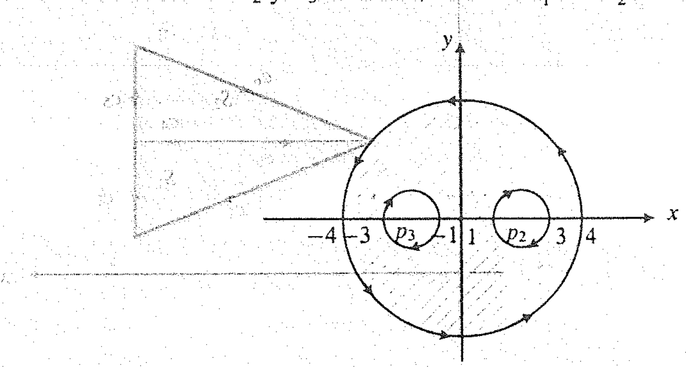


Figura 9. Región S y su frontera positivamente orientada del ejemplo 3.

Los caminos $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3 \colon [0, 2\pi] \to \mathbb{R}^2$ dados por $\boldsymbol{\lambda}_1(t) = (4\cos t, 4 \operatorname{sen} t)$, $\boldsymbol{\lambda}_2(t) = (\cos t + 2, -\operatorname{sen} t)$, $\boldsymbol{\lambda}_3(t) = (\cos t - 2, -\operatorname{sen} t)$ parametrizan ∂R_1^+, ∂R_2^- y ∂R_3^-, respectivamente (es decir, parametrizan ∂S^+). Verifiquemos el teorema de Green con el campo $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (-y, x)$ y la región S. Por una parte se tiene

$$\int_{\partial S^+} F \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}_1} F \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_2} F \cdot d\boldsymbol{\lambda} + \int_{\boldsymbol{\lambda}_3} F \cdot d\boldsymbol{\lambda}$$

$$= \int_0^{2\pi} \big((-4 \operatorname{sen} t)(-4 \operatorname{sen} t) + (4\cos t)(4\cos t)\big)\, dt$$
$$+ \int_0^{2\pi} \big((\operatorname{sen} t)(-\operatorname{sen} t) + (\cos t + 2)(-\cos t)\big)\, dt$$
$$+ \int_0^{2\pi} \big((\operatorname{sen} t)(-\operatorname{sen} t) + (\cos t - 2)(-\cos t)\big)\, dt$$
$$= 32\pi - 2\pi - 2\pi = 28\pi$$

Por otra parte se tiene

$$\int_S\!\!\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dxdy = \int_S\!\!\int (1-(-1))\, dxdy = 2\int_S\!\!\int dxdy$$
$$= 2 \text{ Área de la región } S = 2\big(\text{Área de } R_1 - \text{Área de } R_2 - \text{Área de } R_3\big)$$
$$= 2\big(\pi(4)^2 - \pi(1)^2 - \pi(1)^2\big) = 2(14\pi) = 28\pi$$

lo que comprueba la validez del teorema de Green. ■

Supongamos ahora que la región $S \subseteq \mathbb{R}^2$ es tal que se puede descomponer como una unión finita de regiones de las que hemos venido considerando hasta este momento (que son del tipo I y del tipo II a la vez). En este caso es posible demostrar el teorema de Green, usando la versión que ya ha sido probada. En efecto, consideremos, por ejemplo, la región S mostrada en la figura 10, la cual, no siendo del tipo I y II a la vez (es de hecho del tipo I solamente), se puede descomponer como la unión de dos subregiones S_1 y S_2 las cuales son del tipo I y II. Consideremos las orientaciones positivas de las fronteras de cada una de estas subregiones.

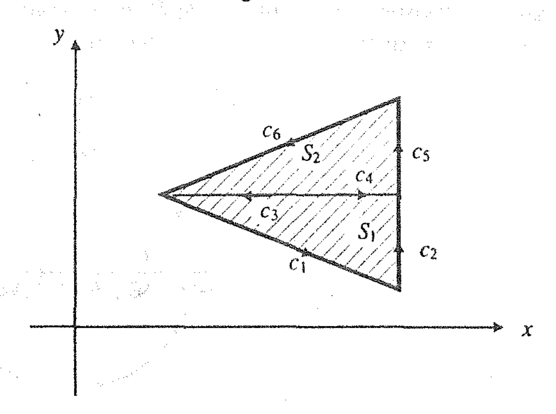


Figura 10. La región S como unión de las suregiones S_1 y S_2 que son del tipo I y II, con sus fronteras positivamente orientadas.

Nótese que el lado que corresponde al corte que hicimos en S para separarlo en S_1 y S_2, el cual es compartido por las correspondientes fronteras de estas subregiones, es recorrido en sentidos contrarios por las orientaciones positivas de estas fronteras. Así, si escribimos $\partial S_1^+ = C_1 \cup C_2 \cup C_3$, $\partial S_2^+ = C_4 \cup C_5 \cup C_6$, como en la figura 10, se tendrá que

$$\int_{C_3} F \cdot d\boldsymbol{\lambda} = -\int_{C_4} F \cdot d\boldsymbol{\lambda}$$

Entonces

$$\iint\limits_{S}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\iint\limits_{S_1}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy+\iint\limits_{S_2}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy$$

$$=\longrightarrow\left(\begin{array}{l}\text{teorema de Green aplicado a las}\\ \text{regiones } S_1 \text{ y } S_2 \text{ (versión ya}\\ \text{demostrada)}\end{array}\right)\longrightarrow=\int_{\partial S_1^+}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{\partial S_2^+}\mathbf{F}\cdot d\boldsymbol{\lambda}$$

$$=\int_{C_1}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_2}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_3}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_4}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_5}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_6}\mathbf{F}\cdot d\boldsymbol{\lambda}$$

$$=\int_{C_1}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_2}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_3}\mathbf{F}\cdot d\boldsymbol{\lambda}-\int_{C_3}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_5}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_6}\mathbf{F}\cdot d\boldsymbol{\lambda}$$

$$=\int_{C_1}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_2}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_5}\mathbf{F}\cdot d\boldsymbol{\lambda}+\int_{C_6}\mathbf{F}\cdot d\boldsymbol{\lambda}=\int_{\partial S^+}F\cdot d\boldsymbol{\lambda}$$

resultado al que queríamos llegar.

Para finalizar esta sección señalamos cómo se puede usar el teorema de Green para calcular áreas de figuras en el plano, digamos, el área de una región compacta S en $\mathbb{R}^2$. Sea $\boldsymbol{\lambda}\colon[a,b]\to\mathbb{R}^2$, $\boldsymbol{\lambda}(t)=(x(t),y(t))$ un camino seccionalmente $\mathscr{C}^1$ que parametriza la frontera de S recorrida positivamente. Tomando por ejemplo el campo $\mathbf{F}\colon\mathbb{R}^2\to\mathbb{R}^2$, $\mathbf{F}(x,y)=(M(x,y),N(x,y))=(0,x)$, para el cual $\frac{\partial M}{\partial y}=0$, $\frac{\partial N}{\partial x}=1$, se obtiene, según el teorema de Green

$$\int_{\partial S_+}\mathbf{F}\cdot d\boldsymbol{\lambda}=\iint\limits_{S}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\iint\limits_{S}dx\,dy=\text{área de } S$$

de modo entonces que (haciendo explícita la integral de línea)

$$\text{área de } S=\int_{\partial S^+}\mathbf{F}\cdot d\boldsymbol{\lambda}=\int_a^b x(t)y'(t)\,dt \tag{A1}$$

De manera análoga, podemos considerar el campo $\mathbf{F}(x,y)=(-y,0)$ para el cual $\frac{\partial M}{\partial y}=-1$, $\frac{\partial N}{\partial x}=0$, que nos da

$$\int_{\partial S^+}\mathbf{F}\cdot d\boldsymbol{\lambda}=\iint\limits_{S}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\iint\limits_{S}dx\,dy=\text{área de } S$$

de donde

$$\text{área de } S=\int_{\partial S^+}\mathbf{F}\cdot d\boldsymbol{\lambda}=-\int_a^b y(t)x'(t)\,dt \tag{A2}$$

Aún más, tomando el campo $\mathbf{F}(x,y)=(-y,x)$ llegamos a

$$\int_{\partial S^+}\mathbf{F}\cdot d\boldsymbol{\lambda}=\iint\limits_{S}\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\iint\limits_{S}(2)\,dx\,dy=2(\text{área de } S)$$

de donde

$$\text{área de } S = \frac{1}{2}\int_{\partial S^+} (-y, x)\cdot d\boldsymbol{\lambda} = \frac{1}{2}\int_a^b (x(t)y'(t) - y(t)x'(t))\,dt \tag{A3}$$

Las fórmulas (A1), (A2) y (A3) nos dan entonces alternativas para calcular el área de una región compacta S de $\mathbb{R}^2$ en términos de integrales de línea de los campos $\mathbf{F}(x, y) = (0, x)$, $\mathbf{F}(x, y) = (-y, 0)$ y $\mathbf{F}(x, y) = (-y, x)$ respectivamente, a lo largo de la frontera de S orientada positivamente. Estos hechos los explotaremos en el apéndice II de esta sección, donde se probará una desigualdad muy famosa en matemáticas: la desigualdad isoperimétrica.

Ejemplo 4. Calculemos el área encerrada por la curva (la traza del camino) $\boldsymbol{\lambda}: [-2, 2] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t^3 - 4t, t^2 - 4)$. Obsérvese que se trata de una curva cerrada simple que se recorre en sentido horario (ver ejemplo 4 de la sección 2 del capítulo 5). Se tiene entonces (anteponiendo un signo menos en la integral de línea para recorrer la curva en sentido positivo) usando, por ejemplo la fórmula (A1)

$$\text{área de } S = -\int_{\partial S^+} (0, x)\cdot d\boldsymbol{\lambda} = -\int_{-2}^{2} (t^3 - 4t)(t^2 - 4)'\,dt$$

$$= -\int_{-2}^{2} (t^3 - 4t)(2t)\,dt = -\left[\frac{2}{5}t^5 - \frac{8}{3}t^3\right]_{-2}^{2} = \frac{256}{15}$$ ■

Apéndice (I) Una demostración del teorema del cambio de variables en integrales dobles

En este apéndice presentamos un argumento que demuestra la fórmula del cambio de variables en integrales dobles estudiada en la sección 4 del capítulo 6 (teorema 6.4.1).

El resultado que queremos probar es el siguiente.

Teorema Teorema del cambio de variables en integrales dobles. Sea $f: R \subset \mathbb{R}^2 \to \mathbb{R}$ una función continua de las variables x, y definida en la región $R \subseteq \mathbb{R}^2$. Sea $\mathbf{T}: R' \subset \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{T}(u, v) = (x, y) = (\phi(u, v), \psi(u, v))$ una función que manda de manera inyectiva los puntos (u, v) de la región R', en los puntos (x, y) de la región R del plano xy. Suponga que $\mathbf{T}$ es de clase $\mathscr{C}^2$ y que la derivada $\mathbf{T}'(u, v)$ es una matriz inversible para todo $(u, v) \in R'$ (de modo equivalente, que el determinante $\det \mathbf{T}'(u, v)$ es no nulo en todos los puntos (u, v) de R'). Entonces se tiene la fórmula de cambio de variables en integrales dobles

$$\iint_R f(x, y)\,dx\,dy = \iint_{R'} f(\phi(u, v), \psi(u, v)) \left|\frac{\partial(\phi, \psi)}{\partial(u, v)}\right| du\,dv$$ ■

El teorema anterior vale cuando $\mathbf{T}$ es de clase $\mathscr{C}^1$. Así quedó enunciado en el teorema 6.4.1. Sin embargo, el argumento que presentaremos requiere que $\mathbf{T}$ sea de clase $\mathscr{C}^2$, pues ocuparemos en alguna parte de él que las derivadas parciales de las funciones coordenadas de $\mathbf{T}$ sean de clase

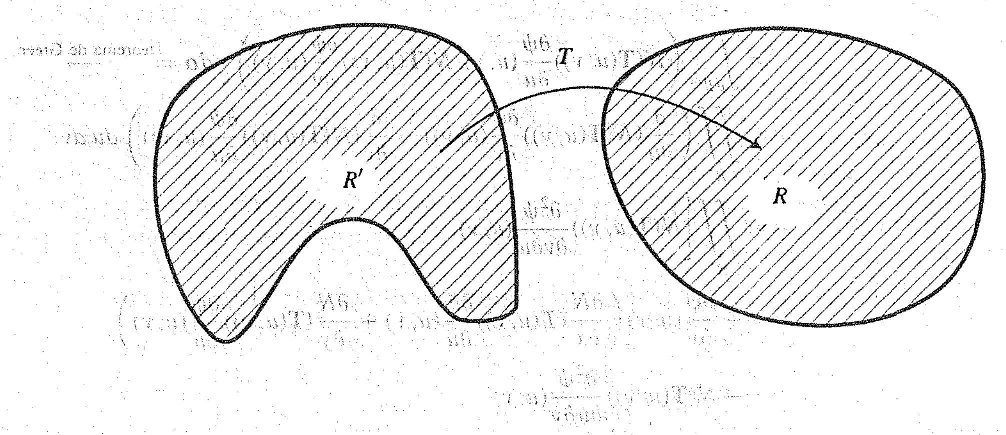


Figura 11. La transformación T manda a la región R' a la región R.

$\mathscr{C}^1$, y que las derivadas parciales cruzadas de estas funciones sean iguales (ver teorema de Schwarz, sección 12 del capítulo 2). Ver esquema figura 11.

Sea $\boldsymbol{\beta}: [a, b] \to \mathbb{R}^2$ un camino cuya imagen es la frontera de la región R orientada positivamente y sea $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = \mathbf{T}^{-1}(\boldsymbol{\beta}(t))$. Supongamos que $\det \mathbf{T}'(u, v) > 0$, $\forall (u, v) \in R'$. Entonces (se puede verificar que, —hágalo) $\boldsymbol{\alpha}$ es un camino cuya imagen es la frontera de la región R' orientada positivamente. Escribamos $\boldsymbol{\beta}(t) = (\beta_1(t), \beta_2(t))$, donde $\beta_1, \beta_2: [a, b] \to \mathbb{R}$ (las funciones coordenadas de $\boldsymbol{\beta}$). Del mismo modo, para $\boldsymbol{\alpha}$, escribimos $\boldsymbol{\alpha}(t) = (\alpha_1(t), \alpha_2(t))$, $t \in [a, b]$. Entonces, como $\boldsymbol{\alpha}(t) = \mathbf{T}^{-1}(\boldsymbol{\beta}(t))$, se tiene $\boldsymbol{\beta}(t) = \mathbf{T}(\boldsymbol{\alpha}(t)) = (\phi(\boldsymbol{\alpha}(t)), \psi(\boldsymbol{\alpha}(t)))$. Es decir, $\beta_1(t) = \phi(\boldsymbol{\alpha}(t))$, $\beta_2(t) = \psi(\boldsymbol{\alpha}(t))$. Consideremos el campo de clase $\mathscr{C}^1$, $\mathbf{F}: R \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (M(x, y), N(x, y)) = (0, N(x, y))$ en donde la función $N: R \subseteq \mathbb{R}^2 \to \mathbb{R}$ es tal que

$$\frac{\partial N}{\partial x}(x, y) = f(x, y), \quad (x, y) \in R$$

Se tiene entonces que

$$\iint_R f(x, y)\, dx\, dy = \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dx\, dy \overset{\text{teorema de Green}}{\longrightarrow}$$

$$= \int_{\partial R^+} \mathbf{F} \cdot d\boldsymbol{\beta} = \int_a^b N(\boldsymbol{\beta}(t))\beta_2'(t)\, dt = \int_a^b N(\boldsymbol{\beta}(t)) \frac{d}{dt}\psi(\boldsymbol{\alpha}(t))\, dt$$

$$= \int_a^b N(\mathbf{T}(\boldsymbol{\alpha}(t))) \left(\frac{\partial \psi}{\partial u}(\boldsymbol{\alpha}(t))\alpha_1'(t) + \frac{\partial \psi}{\partial v}(\boldsymbol{\alpha}(t))\alpha_2'(t) \right) dt$$

$$= \int_a^b \left(N(\mathbf{T}(\boldsymbol{\alpha}(t))) \frac{\partial \psi}{\partial u}(\boldsymbol{\alpha}(t))\alpha_1'(t) + N(\mathbf{T}(\boldsymbol{\alpha}(t))) \frac{\partial \psi}{\partial v}(\boldsymbol{\alpha}(t))\alpha_2'(t) \right) dt$$

[Obsérvese que la expresión anterior es la integral de línea del campo $\mathbf{G}: R' \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{G}(u, v) = \left(N(\mathbf{T}(u, v))\frac{\partial \psi}{\partial u}(u, v), N(\mathbf{T}(u, v))\frac{\partial \psi}{\partial v}(u, v) \right)$ sobre el camino $\boldsymbol{\alpha}$ cuya imagen es $\partial R'^+$. Nótese que este campo es de clase $\mathscr{C}^1$, de modo que podemos aplicar el teorema de Green a la integral de

línea mencionada. Entonces, continuando nuestra cadena de igualdades de la expresión anterior, nos queda]

$$
\begin{aligned}
&= \int_{\partial R'^{+}} \left(N(\mathbf{T}(u, v))\frac{\partial\psi}{\partial u}(u, v), N(\mathbf{T}(u, v))\frac{\partial\psi}{\partial v}(u, v) \right) \cdot d\boldsymbol{\alpha} \overset{\text{teorema de Green}}{=} \\
&= \iint_{R'} \left(\frac{\partial}{\partial u}(N(\mathbf{T}(u, v))\frac{\partial\psi}{\partial v}(u, v)) - \frac{\partial}{\partial v}(N(\mathbf{T}(u, v))\frac{\partial\psi}{\partial u}(u, v)) \right) du\, dv \\
&= \iint_{R'} \left[N(\mathbf{T}(u, v))\frac{\partial^2\psi}{\partial v \partial u}(u, v) \right. \\
&\quad + \frac{\partial\psi}{\partial v}(u, v)\left(\frac{\partial N}{\partial x}(T(u, v))\frac{\partial\phi}{\partial u}(u, v) + \frac{\partial N}{\partial y}(\mathbf{T}(u, v))\frac{\partial\psi}{\partial u}(u, v) \right) \\
&\quad - N(\mathbf{T}(u, v))\frac{\partial^2\psi}{\partial u \partial v}(u, v) \\
&\quad \left. - \frac{\partial\psi}{\partial u}(u, v)\left(\frac{\partial N}{\partial x}(T(u, v))\frac{\partial\phi}{\partial v}(u, v) + \frac{\partial N}{\partial y}(T(u, v))\frac{\partial\psi}{\partial v}(u, v) \right) \right] du\, dv \\
&= \iint_{R'} \frac{\partial N}{\partial x}(\mathbf{T}(u, v))\left(\frac{\partial\phi}{\partial u}(u, v)\frac{\partial\psi}{\partial v}(u, v) - \frac{\partial\phi}{\partial v}(u, v)\frac{\partial\psi}{\partial u}(u, v) \right) du\, dv \\
&= \iint_{R'} f(\phi(u, v), \psi(u, v))\left| \frac{\partial(\phi, \psi)}{\partial(u, v)} \right| du\, dv
\end{aligned}
$$

que es la fórmula a la que queríamos llegar.

Apéndice (II) La desigualdad isoperimétrica

Un problema muy antiguo, conocido desde los griegos, consistía en hallar, de entre todas las figuras planas cerradas con perímetro dado, la que encerraba la mayor área. Se sabe que este problema (entre otros) fue resuelto por Zenodorus (hacia 200 a.C.)[1].

Tal figura es el círculo. Este es un problema que desde entonces ha llamado la atención de los matemáticos, pues, además de que hay muchas maneras de abordarlo, la demostración de que el círculo resuelve el problema involucra siempre detalles no triviales. Uno de estos detalles fue motivo de polémica en el siglo pasado, cuando el matemático alemán Jacob Steiner (1796–1863) presentó una prueba de que el círculo es, efectivamente, la figura cuyo perímetro encierra mayor área. Sin embargo, tal prueba asumía la existencia de la solución del problema, y este es un punto que se debía demostrar. Después de que varios matemáticos trataron de persuadirlo de que su demostración era incompleta, y a quienes contestaba que la existencia de la solución era "evidente", Steiner terminó aceptando su error (prueba de lo cual fue el comentario que incluyó en un artículo en 1842 sobre su demostración, "... y la prueba queda lista si se asume la existencia de la figura"). Fue hasta el año de 1870, cuando el matemático alemán Karl Weierstrass (1815-1897) presentó una prueba satisfactoria

[1]Para ampliar la información histórica de este problema, consultar, por ejemplo, el libro de Morris Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, 1972, pp. 126, 838, 839.

de la existencia de la solución en una conferencia en Berlín. Lo que Weierstrass presentó entonces fueron las ideas involucradas en lo que ahora se conoce como Cálculo de Variaciones: ideas que van en la dirección de maximizar o minimizar determinadas integrales que describen parámetros concretos (tiempo, áreas, etc.). En la actualidad se conocen muchas demostraciones completas de este problema. En esta sección presentamos una de ellas, debida a E. Schmidt (1939), la cual es una demostración de una desigualdad que se conoce como *desigualdad isoperimétrica*, de la que se podrá concluir la solución al problema anteriormente planteado (llamado *problema isoperimétrico*). Una de las herramientas principales involucrada en el argumento de Schmidt es el uso del teorema de Green para calcular áreas de figuras planas, como lo hemos estudiado en esta sección.

El resultado que queremos probar es el siguiente

Teorema (Desigualdad isoperimétrica). Sea C una curva cerrada simple en $\mathbb{R}^2$ cuya longitud es ℓ, y sea A el área de la región encerrada por C. Entonces

$$\ell^2 \geq 4\pi A$$

La igualdad se da si y sólo si C es un círculo.

Demostración. Sean L y L' dos líneas paralelas que no intersectan a la curva C. Acerquemos estas líneas a C (siempre conservando su paralelismo) hasta que cada una de ellas toque a C por primera vez. Llamemos T y T' a las líneas tangentes a C así obtenidas. Coloquemos un sistema coordenado rectangular cuyo eje x sea perpendicular a T y T', y cuyo eje y (que debe ser paralelo a estas rectas) equidiste de T y T'. Sea S un círculo con centro en el origen de este sistema que sea tangente a T (y necesariamente entonces a T'). Se puede tomar el sistema coordenado "lo suficientemente alejado de la curva C" para que el círculo S no corte a C. Llamemos r al radio de S.

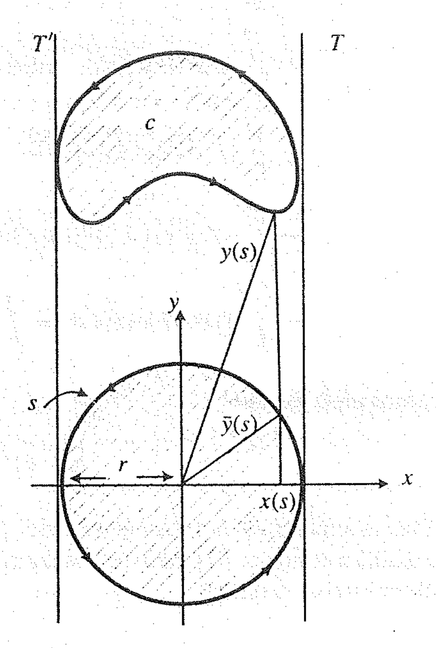


Figura 12. La curva C y el círculo S de la demostración de la desigualdad isoperimétrica.

Consideremos el camino $\boldsymbol{\lambda}: [0, \ell] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(s) = (x(s), y(s))$, que parametriza a la curva C *por longitud de arco* (y entonces $\|\boldsymbol{\lambda}'(s)\| = 1$, $\forall s \in [0, \ell]$). Obsérvese que podemos usar la función $x(s)$ como primera función coordenada de un camino $\boldsymbol{\mu}$ que parametrice al círculo S (de modo que la proyección sobre el eje x del recorrido de $\boldsymbol{\lambda}(s)$ sobre C y $\boldsymbol{\mu}(s)$ sobre S *es la misma*). Sea $\boldsymbol{\mu}: [0, \ell] \to \mathbb{R}^2$, $\boldsymbol{\mu}(s) = (x(s), \tilde{y}(s))$ tal camino (la función $\tilde{y}(s)$ se ha determinado por la expresión $\|\boldsymbol{\mu}(s)\|^2 = (x(s))^2 + (\tilde{y}(s))^2 = r^2$, que obviamente deben satisfacer las coordenadas de $\boldsymbol{\mu}(s)$). Según la fórmula (A1), el área A contenida por la curva C es

$$A = \int_0^\ell x(s)y'(s)\,ds$$

De modo análogo, el área contenida por S es (según la fórmula (A2))

$$\pi r^2 = -\int_0^\ell \tilde{y}(s)x'(s)\,ds$$

Entonces

$$\begin{aligned}
A + \pi r^2 &= \int_0^\ell x(s)y'(s)\,ds - \int_0^\ell \tilde{y}(s)x'(s)\,ds \\
&= \int_0^\ell (x(s)y'(s) - \tilde{y}(s)x'(s))\,ds \\
&\leq \int_0^\ell \sqrt{(x(s)y'(s) - \tilde{y}(s)x'(s))^2}\,ds \\
&= \int_0^\ell \sqrt{x^2(s)y'^2(s) - 2x(s)y'(s)\tilde{y}(s)x'(s) + \tilde{y}^2(s)x'^2(s)}\,ds \\
&\leq \int_0^\ell \sqrt{x^2(s)y'^2(s) + \tilde{y}^2(s)x'^2(s) + x^2(s)x'^2(s) + \tilde{y}^2(s)y'^2(s)}\,ds \\
&= \int_0^\ell \sqrt{(x^2(s) + \tilde{y}^2(s))(x'^2(s) + y'^2(s))}\,ds \\
&= \int_0^\ell \|\boldsymbol{\mu}(s)\|\|\boldsymbol{\lambda}'(s)\|\,ds = \int_0^\ell (r)(1)\,ds = r\ell
\end{aligned}$$

Así pues, hemos probado que

$$A + \pi r^2 \leq r\ell \tag{D}$$

Por otra parte, sabemos que la media geométrica de cualquiera de estos dos números positivos no excede a su media aritmética. Usando este hecho con los números A y πr^2, junto con la desigualdad obtenida anteriormente, tenemos

$$\sqrt{(A)(\pi r^2)} \leq \frac{A + \pi r^2}{2} \leq \frac{r\ell}{2}$$

de donde

$$4(A)(\pi r^2) \leq r^2\ell^2$$

es decir que

$$\ell^2 \geq 4\pi A$$

que es la desigualdad que queríamos establecer. Resta probar que la igualdad en esta desigualdad se tiene si y sólo si la curva C es un círculo.

Es claro que si la curva C es un círculo, se tiene la igualdad en la desigualdad anterior, pues $\ell^2 = (2\pi r)^2 = 4\pi^2 r^2 = 4\pi(\pi r^2) = 4\pi A$. Demostremos que esta igualdad es suficiente para que la curva C sea un círculo. Supongamos entonces que $\ell^2 = 4\pi A$. Queremos probar que el camino $\boldsymbol{\lambda}: [0, \ell] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(s) = (x(s), y(s))$ describe un círculo con centro en el origen (y radio r), es decir, queremos probar que $(x(s))^2 + (y(s))^2 = r^2$. Puesto que $\sqrt{(A)(\pi r^2)} = \frac{r\ell}{2}$ (por hipótesis) y $\sqrt{(A)(\pi r^2)} \leq \frac{A+\pi r^2}{2}$ (la media geométrica de A y πr^2 no excede a su media aritmética), se tiene que $A + \pi r^2 \geq r\ell$, expresión que junto con la desigualdad (D) obtenida anteriormente, nos permite concluir que $A + \pi r^2 = r\ell$. Según la cadena de igualdades y desigualdades con las que concluimos que $A + \pi r^2 \leq r\ell$, se debe tener que (las desigualdades que ahí aparecen *son en realidad igualdades*)

$$x(s)y'(s) - \bar{y}(s)x'(s) = \sqrt{(x(s)y'(s) - \bar{y}(s)x'(s))^2}$$
$$= \sqrt{(x^2(s) + \bar{y}^2(s))(x'^2(s) + y'^2(s))}$$

de donde

$$x(s)x'(s) + \bar{y}(s)y'(s) = 0$$

De aquí, se tiene que

$$x^2(s)x'^2(s) = \bar{y}^2(s)y'^2(s)$$

Sumando en ambos miembros $x^2(s)y'^2(s)$ nos queda

$$x^2(s)(x'^2(s) + y'^2(s)) = y'^2(s)(x^2(s) + \bar{y}^2(s))$$

de donde

$$\frac{x(s)}{y'(s)} = \pm\frac{\sqrt{x^2(s) + \bar{y}^2(s)}}{\sqrt{x'^2(s) + y'^2(s)}} = \pm r$$

Es decir,

$$x(s) = \pm r y'(s) \qquad (*)$$

Llegamos ahora a un punto crucial de la demostración: rehacer todo lo que hemos hecho hasta este momento, con otro sistema coordenado (ver figura 12). Esto, sin embargo, resultará sencillo, pues lo que haremos es *intercambiar los nombres de los ejes*. Ahora el eje "horizontal" será el eje y, y el eje "vertical" será el eje x. Obsérvese que *todo* el argumento anterior funciona perfectamente con este cambio (ahora se tendrá que $\boldsymbol{\lambda}(s) = (y(s), x(s))$, $\boldsymbol{\mu}(s) = (\bar{y}(s), x(s))$, etc.) *simplemente intercambiando las letras* x e y. Concluiremos finalmente entonces que $y(s) = \pm r x'(s)$, expresión análoga a (*) con los nombres de x y y intercambiados. Así entonces concluimos que

$$x^2(s) + y^2(s) = (\pm r y'(s))^2 + (\pm r x'(s))^2 = r^2\|\boldsymbol{\lambda}'(s)\| = r^2$$

Es decir, $\boldsymbol{\lambda}(s) = (x(s), y(s))$, en este caso, un círculo de radio r. ■

Del teorema anterior se desprende entonces que el área A encerrada por una curva cerrada simple de perímetro ℓ no puede exceder el valor $(4\pi)^{-1}\ell^2$, y que este valor se alcanza en el caso (y solamente en el caso) en que la curva sea un círculo. Así pues, de entre todas las curvas cerradas simples de perímetro ℓ, el círculo es la que encierra la mayor área, la cual es $A = (4\pi)^{-1}\ell^2 = (4\pi)^{-1}(2\pi r)^2 = \pi r^2$.

Ejercicios (Capítulo 7, Sección 9)

En cada uno de los ejercicios 1–5, verifique el teorema de Green con el campo y el camino dados.

1. $\mathbf{F}(x, y) = (3x + 2y, x - y)$, $\boldsymbol{\lambda}\colon [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t)$.

2. $\mathbf{F}(x, y) = (x^2 y, y^2 x)$, $\boldsymbol{\lambda}\colon [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (2\cos t, 3 \operatorname{sen} t)$.

3. $\mathbf{F}(x, y) = (x, y)$, $\boldsymbol{\lambda}\colon [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\operatorname{sen}^3 t, \cos^3 t)$.

4. $\mathbf{F}(x, y) = (2xy, 3x^2)$, $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2\colon [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_1(t) = (t, t^2)$, $\boldsymbol{\lambda}_2(t) = ((1 - t)^2, 1 - t)$.

5. $\mathbf{F}(x, y) = (y, x)$, $\boldsymbol{\lambda}$ es un camino cerrado simple.

6. Con el teorema de Green calcule la integral de línea del campo $\mathbf{F}(x, y) = (5x^3 + 4y, 2x - 4y^4)$ a lo largo del círculo $(x - 2)^2 + y^2 = 4$, recorrido en sentido antihorario.

7. Aplique el teorema de Green para calcular la integral de línea del campo $\mathbf{F}(x, y) = (10x^4 - y^3, x^3 - 3y^5)$ a lo largo del círculo $x^2 + y^2 = 5$, recorrido en sentido horario.

8. Con el teorema de Green, calcule la integral de línea del campo $\mathbf{F}(x, y) = (\varphi(x) - y, x + \psi(y))$, donde φ y ψ son dos funciones reales de clase $\mathscr{C}^1$ definidas en $\mathbb{R}$, a lo largo de un cuadrado de lado a recorrido positivamente.

9. Verifique el teorema de Green con el campo $\mathbf{F}(x, y) = (x + y, 2x - y)$, y el camino cuya imagen es la frontera positivamente orientada de la región comprendida entre el círculo $x^2 + y^2 = 9$, y el cuadrado $|x| + |y| = 1$.

10. Verifique el teorema de Green con el campo $\mathbf{F}(x, y) = (3y, 1)$, y el camino cuya imagen es la frontera positivamente orientada de la región interior al círculo $x^2 + y^2 = 25$, y exterior a los cuatro círculos de radio 1 y centros en $(2, 0)$, $(-2, 0)$, $(0, 2)$ y $(0, -2)$.

11. Sea $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ el campo $\mathbf{F}(x, y) = (2xy^2 + y, 2x^2 y + x + x^2/2)$. Demuestre que si $\boldsymbol{\lambda}$ es cualquier camino cerrado simple, cuya imagen es una curva simétrica respecto del eje de ordenadas, entonces la integral de línea de $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ es igual a cero.

12. Considere las integrales

$$I_1 = \int_{\boldsymbol{\lambda}} (2x + y)^2\, dx - (x - 2y)^2\, dy$$

$$I_1 = \int_{\boldsymbol{\mu}} (2x + y)^2\, dx - (x - 2y)^2\, dy$$

donde $\boldsymbol{\lambda}, \boldsymbol{\mu}\colon [0, 1] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (t, t^2)$, $\boldsymbol{\mu}(t) = (t^2, t)$. Tome el teorema de Green para calcular la diferencia $I_1 - I_2$.

En los ejercicios 13–15, aplique el teorema de Green para calcular el área de la figura limitada por la curva dada.

13. la elipse $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (a\cos t, b \operatorname{sen} t)$.

14. la lemniscata de Bernoulli $(x^2+y^2)^2 = a^2(x^2-y^2)$. (Ver ejemplo 8 de la sección 5 del capítulo 6. Una parametrización de la parte de la lemniscata que se encuentra en el primer cuadrante es $\boldsymbol{\lambda}: [0, \pi/4] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (a\cos t\sqrt{\cos t}, a \operatorname{sen} t\sqrt{\cos t})$, la cual se obtiene de la expresión dada haciendo $y = (\tan t)x$).

15. la astroide $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (a\cos^3 t, a \operatorname{sen}^3 t)$.

(*) **16.** Considere la siguiente "demostración" del teorema 7.5.4 en el caso $n = 2$ (si $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ es un campo de clase $\mathscr{C}^k$, $k \geq 1$, definido en el conjunto abierto U de $\mathbb{R}^2$, $\mathbf{F} = (F_1, F_2)$, tal que $\frac{\partial F_2}{\partial x} = \frac{\partial F_1}{\partial y}$ para toda $x \in U$, y si $\boldsymbol{\lambda}, \boldsymbol{\mu}: [a, b] \to U$ son caminos seccionalmente $\mathscr{C}^1$, homotópicos en U —o bien, caminos cerrados seccionalmente $\mathscr{C}^1$, libremente homotópicos en U— entonces $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu}$). Sean $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ como en el enunciado del teorema (ver figura 11). Sea $\mathbf{p} = \boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a)$, $\mathbf{q} = \boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b)$. El camino $\boldsymbol{\rho} = \boldsymbol{\lambda} + (-\boldsymbol{\mu})$ es un camino cerrado seccionalmente $\mathscr{C}^1$ (comienza en $\mathbf{p}$, va sobre $\boldsymbol{\lambda}$ a $\mathbf{q}$, y regresa por $-\boldsymbol{\mu}$ a $\mathbf{p}$) orientado positivamente. Sea $R \subset \mathbb{R}^2$ la región encerrada por el camino $\boldsymbol{\rho}$.
Según el teorema de Green, usando las hipótesis sobre la igualdad de las derivadas parciales de las funciones coordenadas de $\mathbf{F}$, se tiene

$$\int_{\boldsymbol{\rho}} \mathbf{F} \cdot d\boldsymbol{\rho} = \int_{\partial R^+} \mathbf{F} \cdot d\boldsymbol{\rho} = \iint\limits_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dx\, dy = \iint\limits_R 0\, dx\, dy = 0$$

Como

$$\int_{\boldsymbol{\rho}} \mathbf{F} \cdot d\boldsymbol{\rho} = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} + \int_{-\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} - \int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu},$$

se tiene

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} - \int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu} = 0$$

de donde se concluye la igualdad deseada de las integrales de $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ y de $\boldsymbol{\mu}$. Determine dónde se encuentra el error (o los errores) de este razonamiento.

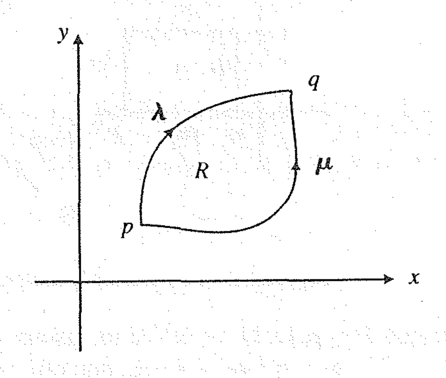


Figura 13. Los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$.

(*) 17. En este ejercicio se considera una demostración del teorema 7.5.4 en el caso $n = 2$, cuando la homotopía $H: [a, b] \times [0, 1] \to \mathbb{R}^2$ entre los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$, es una función de clase $\mathscr{C}^2$. Sea entonces $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (M, N)$, un campo de clase $\mathscr{C}^1$ definido en el conjunto abierto $U \subseteq \mathbb{R}^2$, y sea $\mathbf{T}: V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ una función de clase $\mathscr{C}^2$, definida en el conjunto abierto $V \subseteq \mathbb{R}^2$, cuya imagen se encuentra contenida en U. Escriba $\mathbf{T} = (\varphi, \psi)$, y defina el campo $\mathbf{T}^*\mathbf{F}: V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, como

$$\mathbf{T}^*\mathbf{F} = \left(M \circ \mathbf{T}\frac{\partial \varphi}{\partial x} + N \circ \mathbf{T}\frac{\partial \psi}{\partial x}, M \circ \mathbf{T}\frac{\partial \varphi}{\partial y} + N \circ \mathbf{T}\frac{\partial \psi}{\partial y} \right)$$

a. Demuestre que si $\mathbf{F} = (M, N)$ tiene la propiedad $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$, entonces $\mathbf{T}^*\mathbf{F} = (\alpha, \beta)$ tiene la misma propiedad; es decir, (demuestre que) $\frac{\partial \alpha}{\partial y} = \frac{\partial \beta}{\partial x}$.

b. Sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\lambda}([a, b]) \subseteq V$, $\boldsymbol{\lambda}(t) = (x(t), y(t))$, un camino seccionalmente $\mathscr{C}^1$. La composición $\boldsymbol{\lambda}^* = \mathbf{T} \circ \boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ es un camino seccionalmente $\mathscr{C}^1$ cuya imagen está en U. Demuestre que

$$\int_{\mathbf{T}\circ\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}^* = \int_{\boldsymbol{\lambda}} \mathbf{T}^*\mathbf{F} \cdot d\boldsymbol{\lambda}$$

c. Ponga $\mathbf{T} = H: V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, donde $H: [a, b] \times [0, 1] \to \mathbb{R}^2$ es una homotopia de clase $\mathscr{C}^2$ entre los caminos $\boldsymbol{\lambda}, \boldsymbol{\mu}: [a, b] \to \mathbb{R}^2$ (cuya imagen está contenida en U), siendo V un conjunto abierto que contiene al rectángulo $\Omega = [a, b] \times [0, 1]$. Los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ son seccionalmente $\mathscr{C}^1$, y con las mismas extremidades, es decir $\boldsymbol{\lambda}(a) = \boldsymbol{\mu}(a)$ y $\boldsymbol{\lambda}(b) = \boldsymbol{\mu}(b)$. Sea $\partial\Omega^+$ la frontera positivamente orientada del rectángulo Ω, parametrizada por el camino $\boldsymbol{\rho} = \boldsymbol{\rho}_1 + \boldsymbol{\rho}_2 + \boldsymbol{\rho}_3 + \boldsymbol{\rho}_4$, donde

$$\boldsymbol{\rho}_1: [a, b] \to \mathbb{R}^2, \boldsymbol{\rho}_1(t) = (t, 0)$$
$$\boldsymbol{\rho}_2: [0, 1] \to \mathbb{R}^2, \boldsymbol{\rho}_2(t) = (b, t)$$
$$\boldsymbol{\rho}_3: [a, b] \to \mathbb{R}^2, \boldsymbol{\rho}_3(t) = (a + b - t, 1)$$
$$\boldsymbol{\rho}_4: [0, 1] \to \mathbb{R}^2, \boldsymbol{\rho}_4(t) = (a, 1 - t)$$

Figura 14. La homotopía H entre los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$.

Compruebe que $H \circ \boldsymbol{\rho}: [a, b] \to \mathbb{R}^2$ es un camino cerrado, seccionalmente $\mathscr{C}^1$. Escribiendo $H \circ \boldsymbol{\rho} = H \circ \boldsymbol{\rho}_1 + H \circ \boldsymbol{\rho}_2 + H \circ \boldsymbol{\rho}_3 + H \circ \boldsymbol{\rho}_4$, describa cada uno de los cuatro caminos $H \circ \boldsymbol{\rho}_i$, $i = 1, 2, 3, 4$.

d. Sea $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (M, N)$ un campo de clase $\mathscr{C}^1$, tal que $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Considere el campo $H^*\mathbf{F}: V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$. Use el teorema de Green (y el resultado del inciso a), para concluir que

$$\int_{\partial\Omega_+} H^*\mathbf{F} \cdot d\boldsymbol{\rho} = 0$$

e. Escribiendo

$$\int_{\partial\Omega_+} H^*\mathbf{F} \cdot d\boldsymbol{\rho} = \int_{\rho_1} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_1 + \int_{\rho_2} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_2 + \int_{\rho_3} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_3 + \int_{\rho_4} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_4$$

demuestre, usando el resultado del inciso **b**, que

$$\int_{\rho_2} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_2 = \int_{\rho_4} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_4 = 0,$$

$$\int_{\rho_1} H^*\mathbf{F} \cdot d\boldsymbol{\rho}_1 = \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

$$\int_{\rho_3} H^*F \cdot d\boldsymbol{\rho}_3 = -\int_{\boldsymbol{\mu}} F \cdot d\boldsymbol{\mu}$$

Concluya finalmente que

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\mu}} \mathbf{F} \cdot d\boldsymbol{\mu}$$

Este es el resultado que se quería establecer.

18. Demuestre que no existe curva alguna cerrada simple cuyo perímetro sea de 1 m y que encierre un área de 1 m^2.

19. Existen curvas cerradas simples cuyo perímetro sea (numéricamente) igual al área que encierran?

7.10 Rotación de un campo en $\mathbb{R}^2$

En esta sección exploraremos una nueva interpretación de la integral de línea $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ de un campo continuo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ definido en un conjunto abierto U de $\mathbb{R}^2$ a lo largo de un camino cerrado simple $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ de clase $\mathscr{C}^1$, la cual resulta ser de fundamental importancia en la dinámica de fluidos. Para esto, pensaremos en el campo $\mathbf{F}$ como un *campo de velocidades*, imaginándolo describiendo el flujo de un líquido en una corriente "plana" (en $\mathbb{R}^2$). Así, $\mathbf{F}(x, y) \in \mathbb{R}^2$ nos da el vector velocidad del fluido en el punto $(x, y) \in U$.

Recordemos que la integral de línea del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$ se calcula como

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_a^b \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)\, dt$$

Ésta es pues la integral de los productos punto del campo $\mathbf{F}$ evaluado en los puntos $\boldsymbol{\lambda}(t)$ del camino $\boldsymbol{\lambda}$, con los vectores tangentes al camino $\boldsymbol{\lambda}'(t)$. Sabemos que el producto punto $\mathbf{u} \cdot \mathbf{v}$ entre dos vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ nos da información sobre la proyección del vector $\mathbf{u}$ sobre el vector $\mathbf{v}$. De algún modo

entonces podemos pensar en la integral $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ como la "suma" de las proyecciones del campo de velocidades $\mathbf{F}$ en los puntos $\boldsymbol{\lambda}(t)$ sobre los vectores $\boldsymbol{\lambda}'(t)$ tangentes al camino $\boldsymbol{\lambda}$, es decir, como una suma de las *componentes tangenciales* del campo $\mathbf{F}$ sobre el camino $\boldsymbol{\lambda}$. Estas componentes pueden ser positivas o negativas, según que el vector $\mathbf{F}(\boldsymbol{\lambda}(t))$ forme un ángulo agudo *u* obtuso, respectivamente, con el vector $\boldsymbol{\lambda}'(t)$, lo cual podemos interpretar que el vector $\mathbf{F}(\boldsymbol{\lambda}(t))$, que nos da la velocidad del fluido en $\boldsymbol{\lambda}(t)$, esté "a favor o en contra" (respectivamente) del sentido del recorrido del camino $\boldsymbol{\lambda}$ (el cual es dado por el vector $\boldsymbol{\lambda}'(t)$). Los signos positivo o negativo de las proyecciones $\mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \boldsymbol{\lambda}'(t)$ se reflejarán entonces en el signo de la integral $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$. LLamaremos al valor de esta integral, la *circulación* del campo $\mathbf{F}$ alrededor (o a lo largo) de $\boldsymbol{\lambda}$.

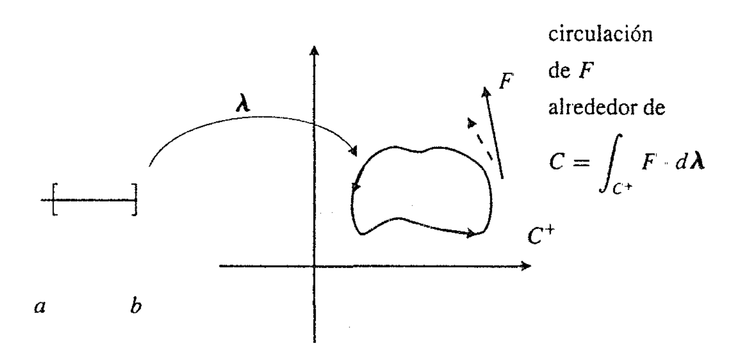


Figura 1. Circulación del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$.

Imaginemos ahora un corcho flotando en una corriente líquida cuyo campo de velocidades está descrito por el campo $\mathbf{F}$. Supongamos que la curva de contacto del corcho con el fluido está descrita por la imagen del camino $\boldsymbol{\lambda}$. Además del movimiento de "arrastre" del corcho por la corriente, ésta puede llegar a provocar un *movimiento de rotación* alrededor de su eje. Este movimiento está de algún modo comprometido con la circulación del campo $\mathbf{F}$ en torno a $\boldsymbol{\lambda}$, como veremos a continuación. Para fijar ideas, supongamos siempre que la curva, imagen del camino $\boldsymbol{\lambda}$ (la curva de contacto del corcho con el líquido), está orientada positivamente. Por simplicidad pensemos por el momento que el corcho es de sección transversal circular. El signo de la circulación $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ nos dirá si el campo $\mathbf{F}$ hace girar al corcho en dirección horaria (si es negativa), en dirección antihoraria (si es positiva) o no lo hace girar (si es cero). Los ejemplos siguientes nos confirmarán estas ideas.

Ejemplo 1. Consideremos el campo $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (1, kx)$, donde k es un número real dado. Sea $\boldsymbol{\lambda}\colon [0, 2\pi] \to \mathbb{R}^2$ el camino $\boldsymbol{\lambda}(t) = (\rho \cos t + a, \rho \operatorname{sen} t + b)$. La curva que describe $\boldsymbol{\lambda}$ es el círculo $(x - a)^2 + (y - b)^2 = \rho^2$, de centro en (a, b) y radio ρ, recorrido positivamente. Calculemos la circulación de $\mathbf{F}$ alrededor de $\boldsymbol{\lambda}$. Se tiene

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_0^{2\pi} \big((-\rho \operatorname{sen} t + k(\rho \cos t + a)(\rho \cos t)\big) dt \\ &= -\rho \int_0^{2\pi} \operatorname{sen} t \, dt + k\rho^2 \int_0^{2\pi} \cos^2 t \, dt + ka\rho \int_0^{2\pi} \cos t \, dt = k\pi\rho^2\end{aligned}$$

Obsérvese que el valor de la integral no depende de las coordenadas a y b del centro del círculo "sobre el que circula el campo $\mathbf{F}$" (lo cual significa que el efecto de la corriente cuyo campo de

velocidades está descrito por **F** es el mismo en cualquier lugar que se coloque el corcho), sino más bien del valor de k. Obsérvese además que el signo de la integral está determinado por el signo de k, teniéndose que:

a. Si $k = 0$, el valor de la circulación del campo **F** alrededor de $\boldsymbol{\lambda}$ es 0. En este caso el campo es $\mathbf{F}(x, y) = (1, 0)$. El corcho que flota sobre la corriente con este campo de velocidades *no gira* sobre su eje.

b. Si $k > 0$, el valor de la circulación del campo **F** alrededor de $\boldsymbol{\lambda}$ es $k\pi\rho^2 > 0$. En este caso el corcho que flota sobre el líquido gira en sentido positivo (antihorario) sobre su eje.

c. Si $k < 0$, el valor de la circulación del campo **F** alrededor de $\boldsymbol{\lambda}$ es $k\pi\rho^2 < 0$. En este caso el corcho que flota sobre el líquido gira en sentido negativo (horario) sobre su eje.

En los casos **b** y **c** las líneas de flujo del campo **F** son las soluciones de la ecuación diferencial $y' = kx$, las cuales son la familia de parábolas $y = 0.5kx^2 + c$.

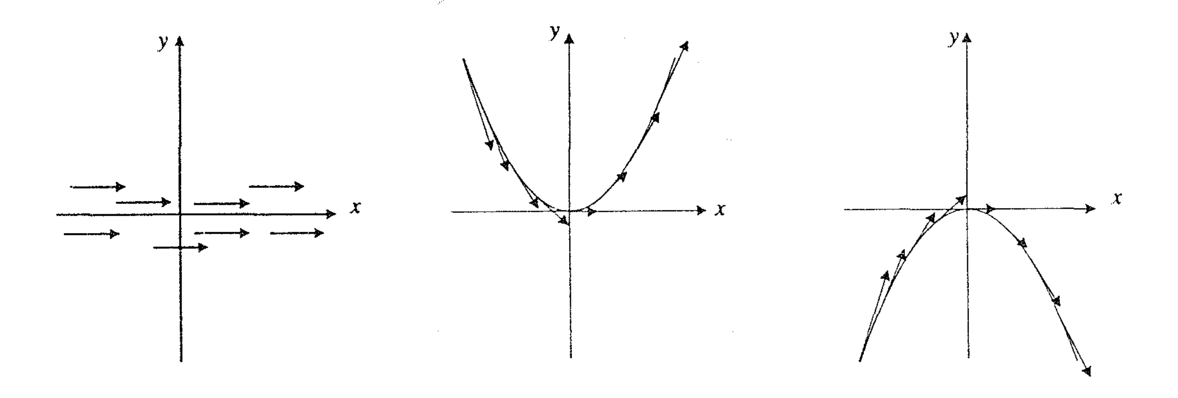


Figura 2. El campo $\mathbf{F}(x, y) = (1, kx)$ con $k = 0, k > 0, k < 0$ ■

Ejemplo 2. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (1, \frac{1}{4}x^2)$ y el camino $\boldsymbol{\lambda}$ del ejemplo anterior $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\rho \cos t + a, \rho \operatorname{sen} t + b)$, (es decir, pongamos el mismo corcho del ejemplo anterior en otra corriente de fluido). La circulación de **F** en torno a $\boldsymbol{\lambda}$ es

$$\begin{aligned}\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} &= \int_0^{2\pi} \left((-\rho \operatorname{sen} t) + \frac{1}{4}(\rho \cos t + a)^2(\rho \cos t)\right) dt \\ &= \frac{1}{2}a\rho^2 \int_0^{2\pi} \cos^2 t \, dt = \frac{1}{2}\pi a \rho^2\end{aligned}$$

Nótese que en este caso el signo de la circulación de **F** en torno a $\boldsymbol{\lambda}$ depende del correspondiente signo de la abscisa del centro del "círculo-corcho" que estamos considerando. Si $a < 0$ (lo cual corresponde a que el corcho se encuentre en la zona del segundo y tercer cuadrantes), entonces la circulación será negativa y la corriente hará girar al corcho sobre su eje en sentido horario. En cambio, si $a > 0$ (el corcho ahora está en la zona del primero y cuarto cuadrantes), la circulación será positiva, y el corcho girará en dirección antihoraria. Las líneas de campo de **F** son las soluciones de

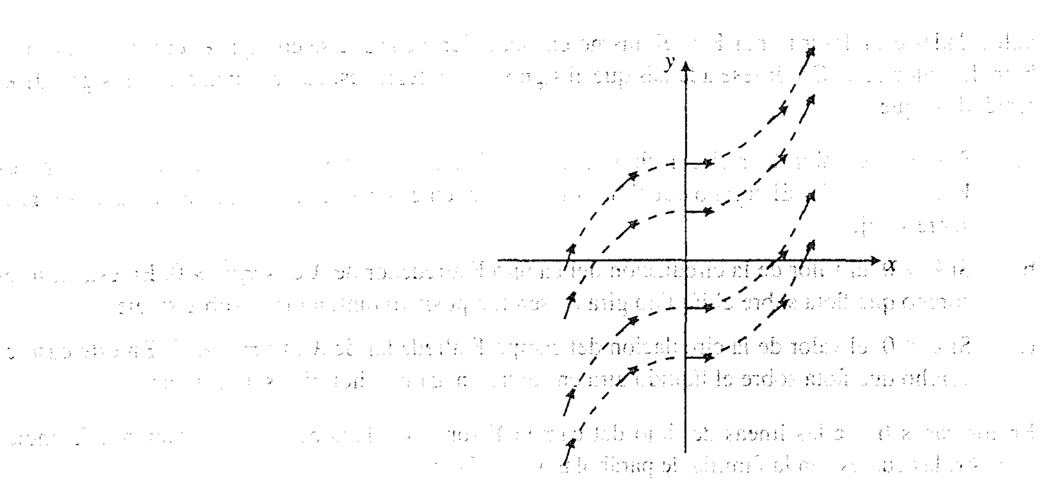


Figura 3. El campo del ejemplo 3.

la ecuación diferencial $y' = \frac{1}{4}x^2$, que son la familia de parábolas cúbicas $y = \frac{1}{12}x^3 + c$. La figura 3 hace referencia a estos resultados.

Ahora nos interesa obtener una estimación de la velocidad angular del corcho que gira en la corriente del líquido cuyo campo de velocidades está descrito por $\mathbf{F}$. Para ello, veamos primeramente un ejemplo muy sencillo que nos dará la pauta para establecer la definición más importante de esta sección. Consideremos el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (-y, x)$ y el camino $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = r(\cos t, \operatorname{sen} t)$. En este caso entonces la curva de contacto del corcho con el líquido es un círculo de centro en el origen y radio $r > 0$, y el campo de velocidades de la corriente en que está el corcho es tal que "todo el campo" contribuye al giro del corcho sobre su eje (los vectores velocidad se encuentran en la dirección tangencial a $\boldsymbol{\lambda}$). Además, en los puntos de la traza de $\boldsymbol{\lambda}$ se tiene $\|\mathbf{F}(x, y)\| = \|(-y, x)\| = \sqrt{y^2 + x^2} = r$, es decir, la rapidez (la magnitud de la velocidad) es constante. Usando la conocida relación de la física elemental $\mathbf{v} = \omega r$, entre la velocidad tangencial $\mathbf{v}$, la velocidad angular ω y el radio r del círculo sobre el que se efectúa el movimiento, concluimos que la velocidad angular del corcho en este caso es constante e igual a 1. Por otra parte, la circulación de $\mathbf{F}$ alrededor de $\boldsymbol{\lambda}$ es

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_0^{2\pi} ((-r \operatorname{sen} t)(-r \operatorname{sen} t) + (r \cos t)(r \cos t))\, dt = 2\pi r^2$$

Podemos escribir entonces que

$$\text{velocidad angular del corcho} = 1 = \frac{1}{2}\left(\frac{1}{\pi\rho^2}\right)(2\pi r^2)$$

$$= \frac{1}{2}\left(\frac{1}{\text{área contenida por } \boldsymbol{\lambda}}\right)(\text{circulación de } \mathbf{F} \text{ en torno a } \boldsymbol{\lambda})$$

Así, el producto de los dos paréntesis del segundo miembro de la última expresión nos da una estimación del doble de la velocidad angular del corcho. Este producto recibe un nombre especial.

Definición. Sea $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ un campo continuo y $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\lambda}([a, b]) \subset U$, una curva cerrada simple de clase $\mathscr{C}^1$. Se llama *rotación* de $\mathbf{F}$ alrededor de $\boldsymbol{\lambda}$, que denotaremos

como rtc $\mathbf{F}(\boldsymbol{\lambda})$, a

$$\begin{aligned}\text{rtc}\,\mathbf{F}(\boldsymbol{\lambda}) &= \left(\frac{1}{\text{área contenida por } \boldsymbol{\lambda}}\right)(\text{circulación de } \mathbf{F} \text{ en torno a } \boldsymbol{\lambda})\\ &= \left(\frac{1}{\text{área contenida por } \boldsymbol{\lambda}}\right)\left(\int_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda}\right)\end{aligned}$$

Así la rotación del campo $\mathbf{F}$ es una medida del doble de la "velocidad angular" del campo alrededor del camino $\boldsymbol{\lambda}$, es decir, del (doble del) número de radianes por segundo que el campo gira alrededor de $\boldsymbol{\lambda}$. Observe además que el signo de la rotación de $\mathbf{F}$ en torno a $\boldsymbol{\lambda}$ está determinado por el signo de su circulación alrededor de $\boldsymbol{\lambda}$, de modo que cuando la rotación es positiva (negativa), la corriente cuyo campo de velocidades está descrito por $\mathbf{F}$ hace girar al corcho en sentido antihorario (horario, respectivamente).

Ejemplo 3. La rotación del campo $\mathbf{F}(x, y) = (1, kx)$ del ejemplo 1 alrededor del camino $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\rho\cos t + a, \rho\,\text{sen}\, t + b)$ (que contiene un área igual $a\pi\rho^2$) es igual a

$$\text{rtc}\,\mathbf{F}(\boldsymbol{\lambda}) = \left(\frac{1}{\text{área contenida por } \boldsymbol{\lambda}}\right)\left(\int_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda}\right) = \frac{1}{\pi\rho^2}k\pi\rho^2 = k \qquad ■$$

Ejemplo 4. La rotación del campo $\mathbf{F}(x, y) = (1, \frac{1}{4}x^2)$ del ejemplo 2 a lo largo del camino $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (r\cos t + a, r\,\text{sen}\, t + b)$ es

$$\text{rtc}\,\mathbf{F}(\boldsymbol{\lambda}) = \left(\frac{1}{\text{área contenida por } \boldsymbol{\lambda}}\right)\left(\int_{\boldsymbol{\lambda}} \mathbf{F}\cdot d\boldsymbol{\lambda}\right) = \frac{1}{\pi\rho^2}\frac{1}{2}\pi a\rho^2 = \frac{a}{2} \qquad ■$$

Aplicando la definición anterior y un proceso al límite, podemos hablar de la propiedad de rotación de un campo *en un punto* dado de su dominio. En efecto, sea $\mathbf{p} = (x_0, y_0)$ un punto dado del abierto U de $\mathbb{R}^2$ en que está definido un campo $\mathbf{F}$. Existe entonces una bola con centro en $\mathbf{p}$ y radio $\epsilon > 0$, $B(\mathbf{p}, \epsilon)$, contenida en U. La frontera de esta bola recorrida positivamente es la traza del camino $\boldsymbol{\lambda}_\epsilon: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_\epsilon(t) = (\epsilon\cos t + x_0, \epsilon\,\text{sen}\, t + y_0)$ que contiene un área de $\pi\epsilon^2$. Definimos la rotación de $\mathbf{F}$ en $\mathbf{p}$, denotada por rtc $F(\mathbf{p})$, como el límite cuando ϵ tiende a cero de las rotaciones de $\mathbf{F}$ alrededor de los caminos $\boldsymbol{\lambda}_\epsilon$. Es decir

$$\text{rtc}\,\mathbf{F}(\mathbf{p}) = \lim_{\epsilon\to 0} \text{rtc}\,\mathbf{F}(\boldsymbol{\lambda}_\epsilon) = \lim_{\epsilon\to 0} \frac{1}{\pi\epsilon^2}\int_{\boldsymbol{\lambda}_\epsilon} \mathbf{F}\cdot d\boldsymbol{\lambda}_\epsilon$$

Ejemplo 5. Sea $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (\alpha xy, \beta x)$. Calculemos la rotación de este campo en un punto cualquiera (x_0, y_0) de $\mathbb{R}^2$. La circulación de $\mathbf{F}$ alrededor de $\boldsymbol{\lambda}_\epsilon: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_\epsilon(t) = (\epsilon\cos t + x_0, \epsilon\,\text{sen}\, t + y_0)$ es

$$\begin{aligned}\int_{\boldsymbol{\lambda}_\epsilon} \mathbf{F}\cdot d\boldsymbol{\lambda}_\epsilon &= \int_0^{2\pi} \big(\alpha(\epsilon\cos t + x_0)(\epsilon\,\text{sen}\, t + y_0)(-\epsilon\,\text{sen}\, t) + \beta(\epsilon\cos t + y_0)(\epsilon\cos t)\big)\, dt\\ &= \int_0^{2\pi} \big(-\alpha\epsilon^3\,\text{sen}^2 t\cos t - \alpha\epsilon^2 y_0\,\text{sen}\, t\cos t - \alpha\epsilon^2 x_0\,\text{sen}^2 t - \alpha\epsilon x_0 y_0\,\text{sen}\, t\\ &\quad + \beta\epsilon^2\cos^2 t + \beta\epsilon y_0\cos t\big)\, dt = \int_0^{2\pi} \big(\beta\epsilon^2\cos^2 t - \alpha\epsilon^2 x_0\,\text{sen}^2 t\big)\, dt\\ &= \beta\epsilon^2\pi - \alpha\epsilon^2 x_0\pi = \pi\epsilon^2(\beta - \alpha x_0)\end{aligned}$$

Entonces la rotación del campo $\mathbf{F}$ alrededor de (x_0, y_0) es

$$\begin{aligned}\operatorname{rtc} \mathbf{F}((x_0, y_0)) &= \lim_{\epsilon \to 0} \operatorname{rtc} \mathbf{F}(\boldsymbol{\lambda}_\epsilon) = \lim_{\epsilon \to 0} \frac{1}{\pi\epsilon^2} \int_{\boldsymbol{\lambda}_\epsilon} \mathbf{F} \cdot d\boldsymbol{\lambda}_\epsilon \\ &= \lim_{\epsilon \to 0} \frac{1}{\pi\epsilon^2} \pi\epsilon^2(\beta - \alpha x_0) = \beta - \alpha x_0\end{aligned}$$

■

Si el campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (M(x, y), N(x, y))$, es de clase $\mathscr{C}^1$, el teorema de Green estudiado en la sección anterior nos permite obtener una expresión muy simple para la rotación del campo en un punto $(x_0, y_0) \in U$. En efecto, se tiene, por el Teorema de Green aplicado a $\mathbf{F}$ y al camino $\boldsymbol{\lambda}_\epsilon$, que

$$\int_{\boldsymbol{\lambda}_\epsilon} \mathbf{F} \cdot d\boldsymbol{\lambda}_\epsilon = \iint_{B_\epsilon} \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dxdy$$

donde B_ϵ es la bola con centro en (x_0, y_0) y radio ϵ. Aplicando el Teorema del Valor Medio (ver sección 5 capítulo 6) para la función continua $\phi(x, y) = \frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y)$ en la región B_ϵ se tiene que

$$\iint_{B_\epsilon} \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dxdy = (\text{área de} B_\epsilon) \left(\frac{\partial N}{\partial x}(\bar{x}, \bar{y}) - \frac{\partial M}{\partial y}(\bar{x}, \bar{y}) \right)$$

para algún punto $(\bar{x}, \bar{y}) \in B_\epsilon$. Como el área de B_ϵ es $\pi\epsilon^2$, se tiene que la circulación de $\mathbf{F}$ en torno a $\boldsymbol{\lambda}_\epsilon$ se puede escribir como

$$\int_{\boldsymbol{\lambda}_\epsilon} \mathbf{F} \cdot d\boldsymbol{\lambda}_\epsilon = \pi\epsilon^2 \left(\frac{\partial N}{\partial x}(\bar{x}, \bar{y}) - \frac{\partial M}{\partial y}(\bar{x}, \bar{y}) \right)$$

Entonces la rotación de $\mathbf{F}$ en (x_0, y_0) es

$$\begin{aligned}\operatorname{rtc} \mathbf{F}((x_0, y_0)) &= \lim_{\epsilon \to 0} \operatorname{rtc} \mathbf{F}(\boldsymbol{\lambda}_\epsilon) = \lim_{\epsilon \to 0} \frac{1}{\pi\epsilon^2} \int_{\boldsymbol{\lambda}_\epsilon} \mathbf{F} \cdot d\boldsymbol{\lambda}_\epsilon \\ &= \lim_{\epsilon \to 0} \frac{1}{\pi\epsilon^2} \pi\epsilon^2 \left(\frac{\partial N}{\partial x}(\bar{x}, \bar{y}) - \frac{\partial M}{\partial y}(\bar{x}, \bar{y}) \right) \\ &= \lim_{\epsilon \to 0} \left(\frac{\partial N}{\partial x}(\bar{x}, \bar{y}) - \frac{\partial M}{\partial y}(\bar{x}, \bar{y}) \right)\end{aligned}$$

Ahora bien, por la continuidad de las funciones $\frac{\partial N}{\partial x}$ y $\frac{\partial M}{\partial y}$, al tender ϵ a cero, el punto $(\bar{x}, \bar{y}) \in B_\epsilon$ debe tender a (x_0, y_0) (¿por qué?), de modo que la rotación de $\mathbf{F}$ en (x_0, y_0) se puede ver como

$$\operatorname{rtc} \mathbf{F}((x_0, y_0)) = \frac{\partial N}{\partial x}(x_0, y_0) - \frac{\partial M}{\partial y}(x_0, y_0)$$

Ejemplo 6. Para el campo $\mathbf{F}(x, y) = (\alpha xy, \beta x)$ del ejemplo 5, se tiene que $\frac{\partial N}{\partial x} = \frac{\partial}{\partial x}(\beta x) = \beta$, $\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}(\alpha xy) = \alpha x$, de modo que la rotación de $\mathbf{F}$ en (x_0, y_0) es

$$\operatorname{rtc} \mathbf{F}((x_0, y_0)) = \frac{\partial N}{\partial x}(x_0, y_0) - \frac{\partial M}{\partial y}(x_0, y_0) = \beta - \alpha x_0$$

como se había obtenido directamente en el ejemplo 5. ■

A un campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ para el que $\operatorname{rtc} \mathbf{F}(x_0, y_0) = 0 \ \forall (x_0, y_0) \in U$, se le llama *campo irrotacional* (en U). Según la fórmula obtenida anteriormente para la rotación del campo $\mathbf{F}$ en un punto, una condición necesaria y suficiente para que el campo $\mathbf{F}$ sea irrotacional en U es que

$$\frac{\partial N}{\partial x}(x_0, y_0) = \frac{\partial M}{\partial y}(x_0, y_0) \quad \forall (x_0, y_0) \in U$$

Si el campo de velocidades de la corriente de un fluido está descrito por un campo irrotacional, significa que "tal corriente no produce movimientos de rotación en un corcho que se coloque en cualquier parte de la corriente". Nótese que si el campo $\mathbf{F}$ es conservativo, entonces es irrotacional, pues en la sección 4 de este capítulo (en el teorema 7.4.2) se estableció la igualdad de las derivadas parciales $\frac{\partial N}{\partial x} = \frac{\partial M}{\partial y}$ como una condición necesaria para que el campo $\mathbf{F}$ sea conservativo (condición equivalente a que $\mathbf{F}$ sea irrotacional). Sin embargo, y de la misma sección 4, podemos concluir que si el campo $\mathbf{F}$ es irrotacional, éste no es necesariamente conservativo. El ejemplo que hemos visto ya de este hecho es el campo $\mathbf{F}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

Este es un campo de clase $\mathscr{C}^\infty$ en el abierto $U = \mathbb{R}^2 - \{(0, 0)\}$. Es irrotacional, puesto que

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}\left(\frac{-y}{x^2 + y^2} \right) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2} \right) = \frac{\partial N}{\partial x}$$

Sin embargo, no es un campo conservativo, como ya se ha visto anteriormente en la sección 4 (ver ejemplo 5 en esa sección). De cualquier manera, no es de esperarse que las propiedades de un campo de ser conservativo e irrotacional sean equivalentes: la primera de ellas es una propiedad global, que depende, entre otras cosas, del dominio en que el campo esté definido, en tanto que la segunda es una propiedad local ("si ponemos un corcho en cualquier *punto* de la corriente, no girará").

Ejemplo 7. Sea $\mathbf{F}_1: \mathbb{R}^2 \to \mathbb{R}^2$, el campo $\mathbf{F}_1(x, y) = (-y, x)$. La rotación de este campo es igual a

$$\operatorname{rtc} \mathbf{F}_1 = \frac{\partial}{\partial x}(x) - \frac{\partial}{\partial y}(-y) = 2$$

Para el campo $\mathbf{F}_2: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$,

$$\mathbf{F}_2(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

su rotación es, como se vio anteriormente, igual a cero. Obsérvese que ambos campos tienen las mismas lineas de flujo, a saber, los círculos concéntricos $x^2 + y^2 = c^2$. Es decir, en ambos campos se tiene

$$(x, y) \cdot \mathbf{F}_i(x, y) = 0, \quad i = 1, 2$$

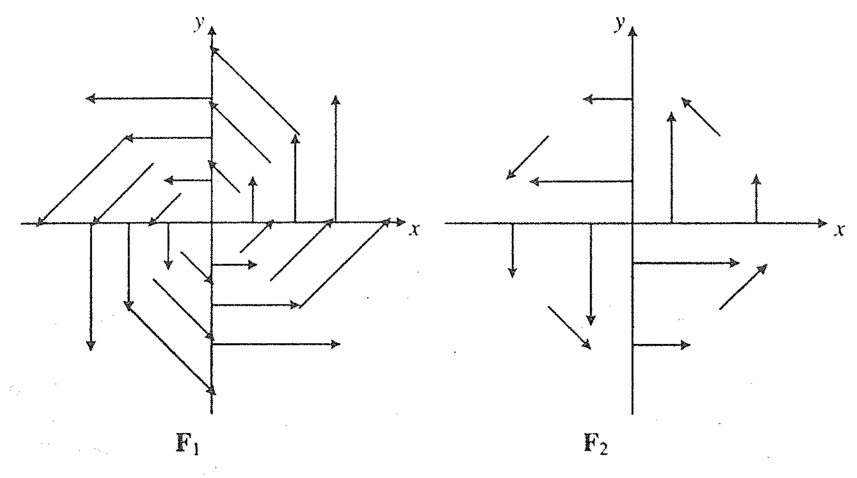


Figura 4. Los campos F_1 y F_2 del ejemplo 7.

Así, ambos campos tienen el mismo "aspecto geométrico", sin embargo F_2 es un campo irrotacional y F_1 no lo es.

Nótese que la magnitud de los vectores $F_1(x, y)$, $\|F_1(x, y)\| = \|(-y, x)\| = \|(x, y)\|$, va aumentando conforme nos retiramos del origen. Es claro entonces que esto contribuirá al efecto de rotación de un corcho que flota en una corriente cuyo campo de velocidades está dado por F_1. En el caso del campo F_2, obsérvese que $\|F_2(x, y)\| = \frac{1}{\|(x,y)\|}$, de modo que la magnitud de los vectores $F_2(x, y)$ va disminuyendo conforme nos alejamos del origen. Esta situación puede explicar el por qué un corcho que flota en una corriente cuyo campo de velocidades es F_2, solamente será "arrastrado" por la corriente (en trayectoria circular), pero no tendrá movimientos rotacionales alrededor de su eje. ■

Ejercicios (Capítulo 7, Sección 10)

En los ejercicios 1–3, calcule la rotación del campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ dado alrededor de la curva indicada.

1. $\mathbf{F}(x, y) = (-y, x)$, alrededor del círculo $x^2 + y^2 = 1$ recorrido en dirección antihoraria.
2. $\mathbf{F}(x, y) = (x^2, xy)$, alrededor del triángulo cuyos vértices son $A = (0, 0)$, $B = (3, 0)$, $C = (2, 2)$, recorrido en dirección antihoraria.
3. $\mathbf{F}(x, y) = (x, 1)$, alrededor del cuadrado $|x| + |y| = 1$, recorrido en dirección horaria.
4. Calcule la rotación del campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x^2 y, xy^2)$ en un punto $\mathbf{p}$ cualquiera.
5. Demuestre que el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (y^2 e^{x+y} + 1, ye^{x+y}(y + 2) + 1)$ es irrotacional.
6. Demuestre que un campo constante $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (a, b)$ es irrotacional. ¿Es conservativo?
7. Demuestre que el campo $\mathbf{F}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2}\right)$$

es irrotacional.

8. Demuestre que la función $f: \mathbb{R}^2 - \{(0,0)\} \to \mathbb{R}$, $f(x, y) = 0.5 \ln(x^2 + y^2)$, es una función potencial del campo del ejercicio anterior. Concluya que el campo es conservativo.

9. Demuestre que el campo $\mathbf{F}: \mathbb{R}^2 - \{(0,0)\} \to \mathbb{R}^2$ dado por

$$\mathbf{F}(x, y) = \frac{e^x}{x^2 + y^2}(x \operatorname{sen} y - y \cos y, x \cos y + y \operatorname{sen} y)$$

es irrotacional.

10. Sea $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t)$. Considere el campo $\mathbf{F}$ del ejercicio anterior. Demuestre que

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_0^{2\pi} e^{\cos t} \cos(\operatorname{sen} t)\, dt$$

Calculando numéricamente esta integral se obtiene el valor 6.2831879. Concluya que el campo $\mathbf{F}$ no es conservativo.

7.11 La divergencia de un campo vectorial (II): campos en $\mathbb{R}^2$

En esta sección introduciremos uno de los conceptos más importantes para campos vectoriales, llamado divergencia, el cual está naturalmente relacionado con la interpretación del campo como un campo de velocidades en una corriente de un fluido. De hecho, el objetivo de esta sección es establecer uno de los resultados célebres del cálculo en $\mathbb{R}^n$, conocido como "Teorema de la Divergencia", en su versión para campos en $\mathbb{R}^2$ (este resultado, en su versión para campos en $\mathbb{R}^3$, requiere el desarrollo de nuevas ideas que estudiaremos en los dos capítulos siguientes).

Supongamos entonces que el campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (M, N)$ definido en el abierto U de $\mathbb{R}^2$ es el campo de velocidades de una corriente de un fluido. Estamos interesados en estimar *cuánto fluido pasa por una "pequeña" porción de U*. Sea $\mathbf{p} = (x_0, y_0) \in U$ y consideremos el rectángulo R con centro en $\mathbf{p}$ dado por

$$R = \{(x, y) | x_0 - h \leq x \leq x_0 + h, y_0 - k \leq y \leq y_0 + k\}$$

Tomando a h y a k suficientemente pequeños podemos asegurar que $R \subset U$.

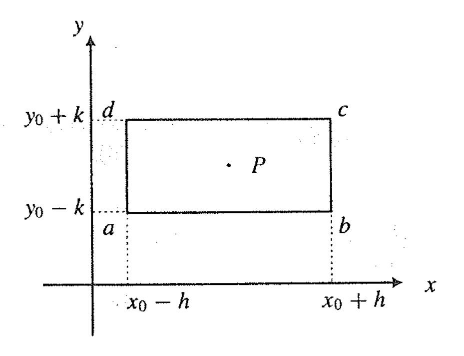


Figura 1. El rectángulo R con centro en $\mathbf{p}$.

De las componentes M y N del campo $\mathbf{F}$, es claro que M es la que "atravesará" los lados ad y bc del rectángulo R, en tanto que N será la que pasará a través de los lados cd y ab. Una estimación de cuanto fluido pasa a través del rectángulo R la podemos obtener "sumando lo que pasa por los lados ab y ad, y restando lo que sale de los lados cd y bc". Tales cantidades se pueden calcular multiplicando a su vez la rapidez con que el fluido atraviesa el lado correspondiente por la longitud del lado. Hagamos el cálculo de la cantidad de fluido que pasa por el lado ab: siendo el lado ab horizontal, la componente del campo $\mathbf{F}$ que atraviesa este lado es N. Si el rectángulo R es pequeño, podemos suponer que *a lo largo de todo el lado* ab *el valor de N es $N(x_0, y_0 - k)$*. Es decir, aproximamos a N, que en general es variable en cada punto del rectángulo, con su valor en el punto medio del lado ab. Multiplicando este valor por $2h$ = longitud del lado ab, logramos una estimación aproximada de la cantidad de fluido que está atravesando ab. Es decir

$$\text{Lo que entra al lado } ab \approx 2hN(x_0, y_0 - k)$$

Consideraciones similares nos conducen a las expresiones

$$\begin{aligned}
\text{Lo que entra al lado } ad &\approx 2kM(x_0 - h, y_0) \\
\text{Lo que sale del lado } cd &\approx 2hN(x_0, y_0 + k) \\
\text{Lo que sale del lado } bc &\approx 2kM(x_0 + h, y_0)
\end{aligned}$$

Entonces la cantidad de fluido que pasa a través del rectángulo R es aproximadamente

$$\begin{aligned}
&\begin{matrix}\text{lo que entra} \\ \text{a } ab\end{matrix} + \begin{matrix}\text{lo que entra} \\ \text{a } ad\end{matrix} - \begin{matrix}\text{lo que sale} \\ \text{de } cd\end{matrix} - \begin{matrix}\text{lo que sale} \\ \text{de } bc\end{matrix} \approx \\
&2hN(x_0, y_0 - k) + 2kM(x_0 - h, y_0) - 2hN(x_0, y_0 + k) - 2kM(x_0 + h, y_0) \\
&\qquad = 2k\big(M(x_0 - h, y_0) - M(x_0 + h, y_0)\big) + 2h\big(N(x_0, y_0 - k) - N(x_0, y_0 + k)\big)
\end{aligned}$$

Dividiendo esta expresión entre el área encerrada por el rectángulo R (igual a 4 hk), obtenemos una medida de la cantidad de fluido que pasa por R *por unidad de área*. Ésta es entonces

$$\frac{M(x_0 - h, y_0) - M(x_0 + h, y_0)}{2h} + \frac{N(x_0, y_0 - k) - N(x_0, y_0 + k)}{2k}$$

Más aún, haciendo que h y k tiendan a cero, obtenemos información sobre cuánto fluido pasa por el punto $\mathbf{p}$ por unidad de área. Ésta es

$$\begin{aligned}
&\lim_{\substack{h \to 0 \\ k \to 0}} \left(\frac{M(x_0 - h, y_0) - M(x_0 + h, y_0)}{2h} + \frac{N(x_0, y_0 - k) - N(x_0, y_0 + k)}{2k} \right) \\
&\qquad = \lim_{h \to 0} \frac{M(x_0 - h, y_0) - M(x_0 + h, y_0)}{2h} + \lim_{k \to 0} \frac{N(x_0, y_0 - k) - N(x_0, y_0 + k)}{2k}
\end{aligned}$$

El lector estará gratamente sorprendido al identificar las expresiones anteriores como las derivadas parciales de M respecto de x y de N respecto de y. Es decir, una estimación de cuánto fluido pasa a través del punto $\mathbf{p}$ está dada por

$$\frac{\partial M}{\partial x}(\mathbf{p}) + \frac{\partial N}{\partial y}(\mathbf{p})$$

Ésta es una cantidad muy importante para el campo $\mathbf{F}$. Se le llama *divergencia de* $\mathbf{F}$ *en* $\mathbf{p}$. En general, se tiene la siguiente

Definición. Sea $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F}(x) = (F_1(x), F_2(x), \ldots, F_n(x))$ un campo diferenciable definido en el abierto U de $\mathbb{R}^n$. Sea $\mathbf{p} = (x_1, x_2, \ldots, x_n) \in U$. Se llama *diveregencia de* $\mathbf{F}$ *en* $\mathbf{p}$, denotada por div $F(\mathbf{p})$, a

$$\operatorname{div} \mathbf{F}(\mathbf{p}) = \frac{\partial F_1}{\partial x_1}(\mathbf{p}) + \frac{\partial F_2}{\partial x_2}(\mathbf{p}) + \cdots + \frac{\partial F_n}{\partial x_n}(\mathbf{p})$$ ■

Nótese que la divergencia de un campo $\mathbf{F}$ es *un escalar*. Es común usar notación de operaciones para escribir la divergencia de un campo. El *operador nabla* ∇ se define como

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)$$

Es decir, ∇ es un vector en $\mathbb{R}^n$ cuya i-ésima coordenada es la derivada parcial (de la función que corresponda a la manera de actuar del operador) respecto de su i-ésima variable. Por ejemplo, si tomamos una función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ (de las estudiadas en el capítulo 2), entonces, para $\mathbf{x} \in U$ se tiene

$$\begin{aligned}\nabla f(\mathbf{x}) &= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right) f(\mathbf{x}) \\ &= \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x})\right) = \operatorname{grad} f(\mathbf{x})\end{aligned}$$

En el caso de la divergencia de un campo $\mathbf{F}$, podemos interpretar la fórmula establecida en la definición anterior como *un producto punto* del operador ∇ con el campo $\mathbf{F}$. En efecto, si $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, se tiene

$$\begin{aligned}\operatorname{div} \mathbf{F}(\mathbf{x}) = \nabla \cdot \mathbf{F} &= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right) \cdot (F_1, F_2, \ldots, F_n) \\ &= \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \cdots + \frac{\partial F_n}{\partial x_n}\end{aligned}$$

Es conveniente tener siempre en mente que la divergencia de un campo $\mathbf{F}$ en un punto $\mathbf{p}$ de su dominio es el límite cuando V tiende a cero del cociente del flujo del campo $\mathbf{F}$ a través de V entre el volumen contenido por V, donde V es cualquier "volumen" que encierra al punto $\mathbf{p}$.

$$\operatorname{div} \mathbf{F}(\mathbf{p}) = \lim_{v \to 0} \frac{\text{flujo del campo a través de } V}{\text{volumen contenido por } V}$$

En el caso que acabamos de estudiar, de un campo en $\mathbb{R}^2$, fue relativamente fácil ver que esta "definición" conduce a la fórmula establecida anteriormente para la divergencia del campo (que nosotros hemos dado como definición). Para campos en $\mathbb{R}^3$ la situación es más complicada y no puede ser hecha en este momento (es necesario, por ejemplo, definir lo que es "flujo" de un campo en $\mathbb{R}^3$), pero advertimos que es posible seguir una discusión similar y llegar a establecer la fórmula de la suma de las derivadas parciales con la que se calcula la divergencia del campo. Insistimos en el "origen" de la fórmula porque en él se puede llegar a tener una idea física del significado de la divergencia de un campo: es una medida del "flujo" (o de la "expansión") por unidad de área (o de volumen, en su caso) del líquido a través del punto $\mathbf{p}$. Para campos en $\mathbb{R}^3$ que modelan el campo de velocidades de un gas en expansión, se puede interpretar la divergencia del campo $\mathbf{F}$ en un

punto **p** como la medida de la velocidad a la que el gas se está "expandiendo" ("está divergiendo") en el punto **p**. Si $\operatorname{div} \mathbf{F}(\mathbf{p}) > 0$, el gas se está "expandiendo" en **p**, si $\operatorname{div} \mathbf{F}(\mathbf{p}) < 0$, el gas se está "comprimiendo" en **p**. En el primer caso, la densidad del gas disminuye en **p** y en el segundo aumenta. A un fluido cuyo campo de velocidades **F** es tal que $\operatorname{div} \mathbf{F}(\mathbf{p}) = 0$ para todo **p**, se le llama *incompresible*.

Ejemplo 1. El campo radial $\mathbf{F}: \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F}(\mathbf{x}) = \mathbf{x}$ tiene por divergencia en un punto $\mathbf{x} \in \mathbb{R}^n$ arbitrario, a

$$\operatorname{div} \mathbf{F}(\mathbf{x}) = \frac{\partial x_1}{\partial x_1} + \frac{\partial x_2}{\partial x_2} + \cdots + \frac{\partial x_n}{\partial x_n} = n$$ ■

Ejemplo 2. El campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (2x + y^2, 3xy + z^3, xyz)$ tiene por divergencia a

$$\begin{aligned} \operatorname{div} \mathbf{F}(x, y, z) &= \frac{\partial}{\partial x}(2x + y^2) + \frac{\partial}{\partial y}(3xy + z^3) + \frac{\partial}{\partial z}(xyz) \\ &= 2 + 3x + xy \end{aligned}$$ ■

Consideremos una región compacta S en $\mathbb{R}^2$ (ver sección 9) y un campo vectorial $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ de clase $\mathscr{C}^1$ definido en el abierto U de $\mathbb{R}^2$ que contiene a S. Supongamos que la frontera de S orientada positivamente ∂S^+ es la imagen del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ seccionalmente $\mathscr{C}^1$. Como sabemos, el vector $T(t) = \frac{\lambda'(t)}{\|\lambda'(t)\|}$ es un vector tangente unitario a ∂S^+. Escribamos este vector como $T(t) = (T_1(t), T_2(t))$. Es fácil ver que el vector $N(t) = (T_2(t), -T_1(t))$ es un vector unitario, perpendicular a $T(t)$, el cual se obtiene al girar este último un ángulo de $\pi/2$ en dirección horaria, es decir, "hacia la derecha de donde está apuntando $T(t)$". Por lo tanto, $N(t)$ es un vector normal unitario a ∂S^+ que "apunta hacia el exterior de S" (recuerde que el interior se encontraba "a la izquierda" de ∂S^+). Llamaremos a $N(t)$ *vector normal unitario exterior a* S.

El producto punto $\mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \mathbf{N}(t)$ nos da información sobre la componente del campo **F** (en los puntos de la frontera de S) en la dirección *exterior* de S. Nuevamente, viendo a S como un campo de velocidades, este producto punto tiene que ver entonces con la "velocidad de flujo hacia el exterior de S". Si integramos la función $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $f(\boldsymbol{\lambda}(t)) = \mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \mathbf{N}(t)$ a lo largo de toda la frontera ∂S^+, obtendremos una estimación de "la velocidad a la que el fluido está saliendo de S". Esta cantidad está íntimamente relacionada con la divergencia del campo en S. Esto es justamente lo que establece el siguiente importante resultado.

Teorema 7.11.1 (El teorema de la divergencia en el plano). Sea S una región compacta en $\mathbb{R}^2$ con su frontera positivamente orientada ∂S^+, la cual es imagen del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ seccionalmente $\mathscr{C}^1$. Sea $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ un campo de clase $\mathscr{C}^1$ definido en el abierto U de $\mathbb{R}^2$ que contiene a S. Si $\mathbf{N}(t)$ es el vector normal unitario exterior a S, se tiene

$$\int_{\partial S^+} \mathbf{F} \cdot \mathbf{N}\, ds = \iint_S \operatorname{div} \mathbf{F}\, dx dy$$

Demostración. Sea $\boldsymbol{\lambda}(t) = (x(t), y(t))$. Entonces los vectores $\mathbf{T}(t)$ y $\mathbf{N}(t)$ son $\mathbf{T}(t) = \frac{1}{\|\lambda'(t)\|}(x'(t), y'(t))$, $\mathbf{N}(t) = \frac{1}{\|\lambda'(t)\|}(y'(t), -x'(t))$. Si $\mathbf{F}(x, y) = (M(x, y), N(x, y))$, entonces la función $\mathbf{F} \cdot \mathbf{N}$ a integrar respecto a la longitud de arco es

$$\mathbf{F}(\boldsymbol{\lambda}(t)) \cdot \mathbf{N}(t) = \frac{1}{\|\boldsymbol{\lambda}'(t)\|}\left(M(\boldsymbol{\lambda}(t))y'(t) - N(\boldsymbol{\lambda}(t))x'(t)\right)$$

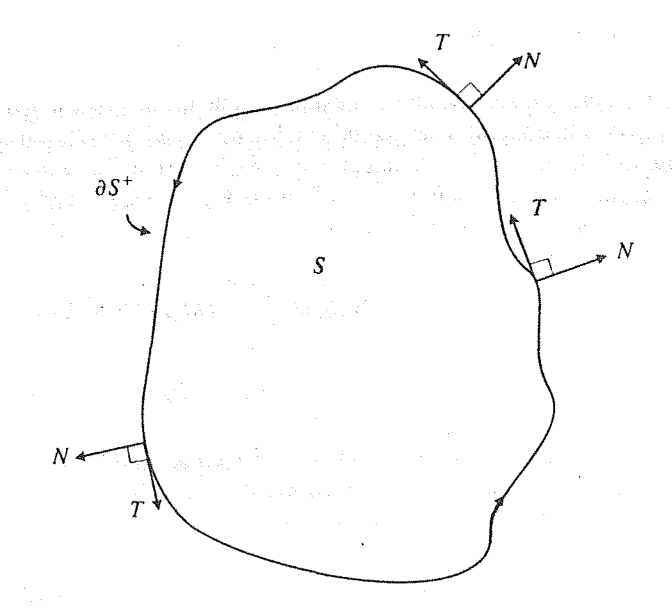


Figura 2. La región S, su frontera ∂S^+ orientada positivamente, y los vectores tangente unitario $T(t)$ y normal (exterior) $N(t)$.

Se tiene entonces que

$$\begin{aligned}\int_{\partial S^+} \mathbf{F}\cdot\mathbf{N}\,ds &= \int_a^b \frac{1}{\|\boldsymbol{\lambda}'(t)\|}\big(M(\boldsymbol{\lambda}(t))y'(t) - N(\boldsymbol{\lambda}(t))x'(t)\big)\|\boldsymbol{\lambda}'(t)\| \\ &= \int_a^b \big(M(\boldsymbol{\lambda}(t))y'(t) - N(\boldsymbol{\lambda}(t))x'(t)\big)\,dt\end{aligned}$$

Consideremos el campo $G\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $G(x, y) = (-N(x, y), M(x, y))$. Nótese que

$$\int_{\partial S^+} \mathbf{G}\cdot d\boldsymbol{\lambda} = \int_a^b \big(-N(\boldsymbol{\lambda}(t))x'(t) + M(\boldsymbol{\lambda}(t))y'(t)\big)\,dt = \int_{\partial S^+} \mathbf{F}\cdot\mathbf{N}\,ds$$

Por otra parte, usando el Teorema de Green con la integral de línea del campo $\mathbf{G}$ a lo largo de $\boldsymbol{\lambda}$ se tiene

$$\begin{aligned}\int_{\partial S^+} \mathbf{G}\cdot d\boldsymbol{\lambda} &= \iint_S \left(\frac{\partial M}{\partial x} - \frac{\partial(-N)}{\partial y}\right)dxdy = \iint_S \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right)dxdy \\ &= \iint_S \operatorname{div}\mathbf{F}\,dxdy\end{aligned}$$

Es decir

$$\int_{\partial S^+} \mathbf{F}\cdot\mathbf{N}\,ds = \iint_S \int \operatorname{div}\mathbf{F}\,dxdy$$

como queríamos probar. Q.E.D.

Así pues, el resultado anterior establece que la velocidad a la que el fluido sale por la frontera de S es igual a la integral (a la "suma") de las velocidades de "expansión" del fluido por unidad de área sobre toda la región S. A la integral $\int_{\partial S^{+}} \mathbf{F}\cdot\mathbf{N}\,ds$ se le llama *flujo de* $\mathbf{F}$ *a través de* S. Nótese que si S es el rectángulo de la figura 1, el flujo de $\mathbf{F}$ a través, por ejemplo, del lado ab, parametrizado por $\boldsymbol{\lambda}_{ab}\colon [x_0 - h, x_0 + h] \to \mathbb{R}^2$, $\boldsymbol{\lambda}_{ab}(t) = (t, y_0 - k)$, para el cual un vector normal unitario exterior es $\mathbf{N} = (0, -1)$ es

$$\int_{\partial S^{+}} \mathbf{F}\cdot\mathbf{N}\,ds = \int_{x_0-h}^{x_0+h} (M(\boldsymbol{\lambda}(t)), N(\boldsymbol{\lambda}(t)))\cdot(0,-1)\,dt$$

$$= -\int_{x_0-h}^{x_0+h} \mathrm{N}(\boldsymbol{\lambda}(t))\,dt$$

Aplicando el teorema del valor medio a esta integral, suponiendo que el punto en el que se verifica tal teorema es el punto medio del intervalo de intergración (es decir, el punto correspondiente a $(x_0, y_0 - k)$), nos queda la siguiente aproximación

$$\int_{\partial S^{+}} \mathbf{F}\cdot\mathbf{N}\,ds \approx -N(x_0, y_0 - k)(2h) = -2hN(x_0, y_0 - k)$$

expresión que ya habíamos usado para el flujo de $\mathbf{F}$ a través del lado ab del rectángulo R (con un signo menos al principio, ¿por qué?).

Ejemplo 3. Verifiquemos el teorema de la divergencia con el campo radial $\mathbf{F}\colon\mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (x, y)$, y $S = \{(x, y) | x^2 + y^2 \le \rho^2\}$. Un camino que parametriza la frontera positivamente orientada de S es $\boldsymbol{\lambda}\colon [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = \rho(\cos t, \operatorname{sen} t)$. Observe que (por ser el campo radial), normalizando a $\mathbf{F}(\boldsymbol{\lambda}(t))$ obtenemos el vector normal unitario exterior a S. Este es entonces $N(t) = (\cos t, \operatorname{sen} t)$. Se tiene, por una parte, que

$$\int_{\partial S^{+}} \mathbf{F}\cdot\mathbf{N}\,ds = \int_0^{2\pi} \mathbf{F}(\boldsymbol{\lambda}(t))\cdot\mathbf{N}(t)\|\boldsymbol{\lambda}'(t)\| = \int_0^{2\pi} \rho^2\,dt = 2\pi\rho^2$$

Por otra parte, la divergencia de $\mathbf{F}$ es constante e igual a 2 (ver ejemplo 1). Entonces

$$\iint_S \operatorname{div}\mathbf{F}\,dx\,dy = 2\iint_S dx\,dy = 2(\text{área de } S) = 2\pi\rho^2 = \int_{\partial S^{+}} \mathbf{F}\cdot\mathbf{N}\,ds$$ ■

Ejemplo 4. Consideremos el campo $\mathbf{F}\colon\mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (-x, -y)$. Este es un campo radial con las flechas apuntando hacia el origen. Sea $S = \{(x, y) | x + y \le 1, x \ge 0, y \ge 0\}$. Calculemos el flujo de $\mathbf{F}$ a través de S.

Un camino $\boldsymbol{\lambda}$ seccionalmente $\mathscr{C}^1$ cuya imagen es la frontera positivamente orientada de S es $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3$, donde

$$\boldsymbol{\lambda}_1\colon [0,1] \to \mathbb{R}^2, \quad \boldsymbol{\lambda}_1(t) = (t, 0)$$
$$\boldsymbol{\lambda}_2\colon [0,1] \to \mathbb{R}^2, \quad \boldsymbol{\lambda}_2(t) = (1 - t, t)$$
$$\boldsymbol{\lambda}_3\colon [0,1] \to \mathbb{R}^2, \quad \boldsymbol{\lambda}_3(t) = (0, 1 - t)$$

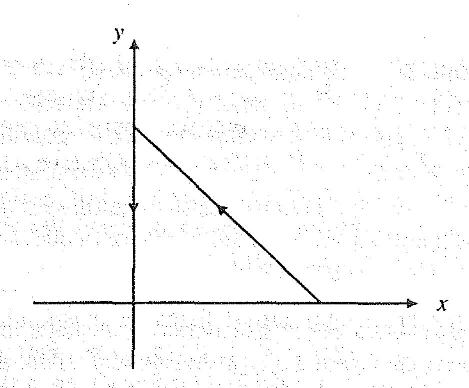


Figura 3. Región S del ejemplo 4.

Llamemos C_i a la curva imagen del camino $\boldsymbol{\lambda}_i$, $i = 1, 2, 3$. Calculemos el flujo de $\mathbf{F}$ a través de C_1. Un vector normal unitario exterior a C_1 es $\mathbf{N} = (0, -1)$. Entonces

$$\int_{C_1} \mathbf{F} \cdot \mathbf{N}\, ds = \int_0^1 (-t, 0) \cdot (0, -1) \|(1, 0)\|\, dt = 0$$

Viendo el campo $\mathbf{F}$ es natural que la integral anterior sea 0: el campo es tangente a C_1, no "atraviesa" esta parte de la frontera de S. Esta situación ocurre también con C_3. Dejamos al lector que haga los cálculos correspondientes para que se convenza de que

$$\int_{C_3} \mathbf{F} \cdot \mathbf{N}\, ds = 0$$

Calculemos entonces el flujo de $\mathbf{F}$ a través de C_2. Un vector normal unitario exterior a C_2 es $\mathbf{N} = \frac{1}{\sqrt{2}}(1, 1)$. Entonces

$$\begin{aligned}\int_{C_2} \mathbf{F} \cdot \mathbf{N}\, ds &= \int_0^1 (-(1-t), -t) \cdot \frac{1}{\sqrt{2}}(1, 1) \|\boldsymbol{\lambda}_2'(t)\|\, dt \\ &= \int_0^1 (-1)\, dt = -1\end{aligned}$$

Entonces el flujo a través de S es

$$\int_{\partial S^+} \mathbf{F} \cdot \mathbf{N}\, ds = \int_{C_1} \mathbf{F} \cdot \mathbf{N}\, ds + \int_{C_2} \mathbf{F} \cdot \mathbf{N}\, ds + \int_{C_3} \mathbf{F} \cdot \mathbf{N}\, ds = 0 - 1 + 0 = -1$$

Llamamos la atención al signo del resultado obtenido. Éste significa que "el flujo es hacia el interior de S", como se puede fácilmente apreciar de la figura 3. Verifiquemos este resultado con el teorema de la divergencia estudiado anteriormente. Se tiene en este caso que $\operatorname{div} \mathbf{F} = -2$. Entonces

$$\begin{aligned}\iint_S \operatorname{div} \mathbf{F}\, dx\, dy &= \iint_S (-2)\, dx\, dy = -2 \iint_S dx\, dy = -2(\text{area de } S) \\ &= -2\left(\frac{1}{2}\right) = -1\end{aligned}$$

■

Ejemplo 5. El flujo del campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(x, y) = (-y, x)$ a través de cualquier región compacta S en $\mathbb{R}^2$ es igual a cero. En efecto, la divergencia de este campo es $\operatorname{div} \mathbf{F}(x, y) = \frac{\partial}{\partial x}(-y) + \frac{\partial}{\partial y}(x) = 0$, y entonces, por el teorema de la divergencia se tiene que $\iint_{\partial S_+} \mathbf{F} \cdot N\, ds = \iint_S \operatorname{div} \mathbf{F}\, dx\, dy = \iint_S 0\, dx\, dy = 0$. Observe que las líneas de flujo de este campo son los círculos $x^2 + y^2 = c^2$. Si S es uno de estos círculos, es claro que el flujo "no entra" a S (se mantiene tangente). En cualquier otro caso, lo que acontece es que la velocidad a la que el flujo entra a S se equilibra con la velocidad a la que sale de S. ■

Quisiéramos por último llamar la atención hacia la "simetría" que presenta el contenido de los teoremas de Green y de la divergencia. Para poder apreciarla, debemos escribir de una manera adecuada la fórmula del Teorema de Green, a saber

$$\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} = \iint_S \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dxdy$$

Se había observado en la sección 7 que la integral de línea del campo de clase $\mathscr{C}^1$, $\mathbf{F} = (M, N)$ a lo largo de $\boldsymbol{\lambda}$ (el camino seccionalmente $\mathscr{C}^1$ que parametriza a ∂S^+) no era más que la integral repecto de la longitud de arco de la componente ortogonal de $\mathbf{F}$ en la dirección del vector tangente unitario $\mathbf{T}(t)$ (el vector $\boldsymbol{\lambda}'(t)$ normalizado). Entonces podemos escribir la primera integral de la fórmula anterior como $\int_{\partial S^+} \mathbf{F} \cdot \mathbf{T}\, ds$. Usando ahora el concepto de rotación del campo $\mathbf{F}$ estudiado en la sección anterior, podemos escribir también la integral del lado derecho como $\iint_S \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dx\, dy = \iint_S \operatorname{rtc} \mathbf{F}\, dxdy$. De esta manera, el Teorema de Green establece que

$$\int_{\partial S^+} \mathbf{F} \cdot \mathbf{T}\, ds = \iint_S \operatorname{rtc} \mathbf{F}\, dxdy$$

Escribamos ahora juntas las fórmulas de los teoremas de Green y de la divergencia

Teorema de Green: $$\int_{\partial S^+} \mathbf{F} \cdot \mathbf{T}\, ds = \iint_S \operatorname{rtc} \mathbf{F}\, dxdy$$

Teorema de la divergencia: $$\int_{\partial S^+} \mathbf{F} \cdot \mathbf{N}\, ds = \iint_S \operatorname{div} \mathbf{F}\, dxdy$$

Así pues, lo que establecen estos resultados son las relaciones entre las integrales de las componentes tangencial y normal del campo sobre la curva ∂S^+ con integrales dobles sobre el interior de la curva. El teorema de Green dice que la integral de la componente tangencial del campo a lo largo de la curva es igual a la "suma de los efectos rotacionales" del campo en el interior, y el teorema de la divergencia dice que la integral de la componente normal del campo sobre la curva es igual a la "suma de los efectos divergentes" del campo en el interior.

Apéndice La divergencia en los sistemas de coordenadas cilíndricas y esféricas

En este apéndice se establecerán las expresiones para la divergencia de un campo cuando éste está dado en el sistema coordenado cilíndrico y esférico. Para el caso del sistema cilíndrico, aprovechando

que éste no es más que el sistema polar del plano con la tercera coordenada (cartesiana) igual, haremos una discusión semejante a la presentada en esta sección para obtener la expresión de la divergencia en un punto **p**, viéndola como el límite del flujo del campo a través de una pequeña región (alrededor del punto **p**) dividida entre el área de la región, cuando ésta tiende a cero.

Sea entonces **p** el punto de coordenadas (r_0, θ_0) en el sistema polar de $\mathbb{R}^2$. Con líneas del tipo $r =$ cte. y $\theta =$ cte. construyamos la región R que circunda a **p** siguiente:

$$R = \{(r, \theta) | r_0 - h \leq r \leq r_0 + h, \theta_0 - k \leq \theta \leq \theta_0 + k\}$$

Esta región corresponde a un "rectángulo" con centro en **p**, cuyos lados son las líneas radiales $\theta = \theta_0 - k$, $\theta = \theta_0 + k$, y los círculos $r = r_0 - h$, $r = r_0 + h$, como se muestra en la figura 4.

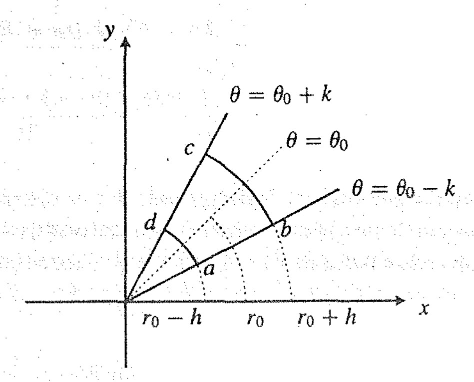


Figura 4. La región R alrededor del punto p en el sistema polar.

El campo **F** se escribe en el sistema polar como $\mathbf{F}(r, \theta) = (F_r(r, \theta), F_\theta(r, \theta))$, en el que F_r y F_θ son las funciones coordenadas de **F**, o bien, son las proyecciones sobre los vectores de la base ortonormal $\{e_r, e_\theta\}$ del sistema polar en $\mathbb{R}^2$ (ver ejemplo 3 de la sección 4, capítulo 1). El flujo a través del rectángulo R lo medimos, como en el caso cartesiano, como la suma del flujo que entra a los lados ad y ab, menos lo que sale de los lados bc y cd. Es claro que la componente de **F** responsable de atravesar los lados ab y cd es F_θ, en tanto que la que atraviesa los lados ad y bc es F_r. Se tiene:

$$\begin{aligned} \text{flujo a través de } ab &\approx (\text{Valor de } F_\theta \text{ en el punto medio de } ab) \text{ (longitud de } ab) \\ &= 2hF_\theta(r_0, \theta_0 - k) \end{aligned}$$

En forma análoga:

$$\begin{aligned} \text{flujo a través de } bc &\approx (\text{Valor de } F_r \text{ en el punto medio de } bc) \text{ (longitud de } bc) \\ &= 2kF_r(r_0 + h, \theta_0) \end{aligned}$$

(en este caso el lado bc es una parte del perímetro de un círculo de radio $r_0 + h$, correspondiente al ángulo central $2k$, por lo que su longitud es $2k(r_0 + h)$). De la misma manera se obtiene que

$$\text{flujo a través de } cd \approx 2hF_\theta(r_0, \theta_0 + k)$$
$$\text{flujo a través de } ad \approx 2k(r_0 - h)F_r(r_0 - h, \theta_0)$$

Entonces el flujo a través de R es aproximadamente igual a

$$\begin{aligned} 2hF_\theta(r_0, \theta_0 - k) \quad &+ 2k(r_0 + h)F_r(r_0 + h, \theta_0) \\ &- 2hF_\theta(r_0, \theta_0 + k) - 2k(r_0 - h)F_r(r_0 - h, \theta_0) \\ &= 2k\big((r_0 + h)F_r(r_0 + h, \theta_0) - (r_0 - h)F_r(r_0 - h, \theta_0)\big) \\ &+ 2h\big(F_\theta(r_0, \theta_0 - k) - F_\theta(r_0, \theta_0 + k)\big) \end{aligned}$$

El área del rectángulo R es $4hkr_0$. Dividiendo la expresión anterior entre este valor obtenemos el flujo a través de R por unidad de área, el cual es entonces igual a

$$\begin{aligned} &= \frac{1}{r_0}\frac{(r_0 + h)F_r(r_0 + h, \theta_0) - (r_0 - h)F_r(r_0 - h, \theta_0)}{2h} \\ &+ \frac{1}{r_0}\frac{F_\theta(r_0, \theta_0 - k) - F_\theta(r_0, \theta_0 + k)}{2k} \end{aligned}$$

Tomando por último límite cuando h y k tienden a cero en la expresión anterior, identificamos fácilmente en el primer sumando la derivada parcial del producto rF_r respecto de r, y en el segundo la derivada parcial de F_θ respecto de θ. Este límite es justamente la divergencia de $\mathbf{F}$ en $\mathbf{p}$. Es decir, la expresión de la divergencia del campo $\mathbf{F} = (F_r, F_\theta)$ en el punto $\mathbf{p}$ de coordenadas (r, θ) es

$$\operatorname{div}\mathbf{F} = \frac{1}{r}\frac{\partial}{\partial r}(rF_r) + \frac{1}{r}\frac{\partial F_\theta}{\partial \theta}$$

No será ahora difícil de aceptar que la expresión de la divergencia del campo $\mathbf{F} = (F_r, F_\theta, F_z)$ en coordenadas cilíndricas es

$$\boxed{\operatorname{div}\mathbf{F} = \frac{1}{r}\frac{\partial}{\partial r}(rF_r) + \frac{1}{r}\frac{\partial F_\theta}{\partial \theta} + \frac{\partial F_z}{\partial z}}$$

Haciendo uso de la expresión explícita del campo $\mathbf{F} = (F_r, F_\theta, F_z)$ en coordenadas cilíndricas en términos de la correspondiente expresión del campo en el sistema cartesiano $\mathbf{F} = (F_x, F_y, F_z)$, es posible verificar que efectivamente el miembro derecho de la expresión obtenida anteriormente cooresponde a la expresión ya conocida de la divergencia de $\mathbf{F}$ (dada como definición). A manera de ejercicio, hagamos estos cálculos. En el apéndice de la sección 2 se obtuvo que la expresión del campo $\mathbf{F}$ en coordenadas cilíndricas es

$$\begin{aligned} \mathbf{F}(r\cos\theta, r\operatorname{sen}\theta, z) = \tilde{F}(r, \theta, z) &= (F_r(r, \theta, z), F_\theta(r, \theta, z), F_z(r, \theta.z)) \\ &= F_r(r, \theta, z)\mathbf{e}_r + F_\theta(r, \theta, z)\mathbf{e}_\theta + F_z(r, \theta, z)\mathbf{e}_z \end{aligned}$$

donde

$$\begin{aligned} F_r(r, \theta, z) &= \cos\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z) + \operatorname{sen}\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z) \\ F_\theta(r, \theta, z) &= -\operatorname{sen}\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z) + \cos\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z) \\ F_z(r, \theta, z) &= F_z(r\cos\theta, r\operatorname{sen}\theta, z) \end{aligned}$$

Entonces

$$\begin{aligned}\frac{\partial}{\partial r}(rF_r) &= \frac{\partial}{\partial r}\big(r\cos\theta F_x(r\cos\theta, r\,\mathrm{sen}\,\theta, z) + r\,\mathrm{sen}\,\theta F_y(r\cos\theta, r\,\mathrm{sen}\,\theta, z)\big)\\ &= r\cos\theta\left(\cos\theta\frac{\partial F_x}{\partial x} + \mathrm{sen}\,\theta\frac{\partial F_x}{\partial y}\right) + \cos\theta F_x\\ &\quad + r\,\mathrm{sen}\,\theta\left(\cos\theta\frac{\partial F_y}{\partial x} + \mathrm{sen}\,\theta\frac{\partial F_y}{\partial y}\right) + \mathrm{sen}\,\theta F_y\\ &= r\cos^2\theta\frac{\partial F_x}{\partial x} + r\,\mathrm{sen}\,\theta\cos\theta\left(\frac{\partial F_x}{\partial y} + \frac{\partial F_y}{\partial x}\right) + r\,\mathrm{sen}^2\,\theta\frac{\partial F_y}{\partial y}\\ &\quad + \cos\theta F_x + \mathrm{sen}\,\theta F_y\end{aligned}$$

en que las funciones F_x, F_y y sus derivadas parciales están evaluadas en $(r\cos\theta, r\,\mathrm{sen}\,\theta, z)$. Por otra parte

$$\begin{aligned}\frac{\partial F_\theta}{\partial\theta} &= \frac{\partial}{\partial\theta}\big(-\mathrm{sen}\,\theta F_x(r\cos\theta, r\,\mathrm{sen}\,\theta, z) + \cos\theta F_y(r\cos\theta, r\,\mathrm{sen}\,\theta, z)\big)\\ &= -\,\mathrm{sen}\,\theta\left(-r\,\mathrm{sen}\,\theta\frac{\partial F_x}{\partial x} + r\cos\theta\frac{\partial F_x}{\partial y}\right) - \cos\theta F_x\\ &\quad + \cos\theta\left(-r\,\mathrm{sen}\,\theta\frac{\partial F_y}{\partial x} + r\cos\theta\frac{\partial F_y}{\partial y}\right) - \mathrm{sen}\,\theta F_y\\ &= r\,\mathrm{sen}^2\,\theta\frac{\partial F_x}{\partial x} - r\,\mathrm{sen}\,\theta\cos\theta\left(\frac{\partial F_x}{\partial y} + \frac{\partial F_y}{\partial x}\right) + r\cos^2\theta\frac{\partial F_y}{\partial y}\\ &\quad - \cos\theta F_x - \mathrm{sen}\,\theta F_y\end{aligned}$$

Es claro que la derivada parcial de F_z (en cilídricas) respecto de z es la misma que la de F_z (en cartesianas) respecto de z. Entonces se tiene que

$$\begin{aligned}&\frac{1}{r}\frac{\partial}{\partial r}(rF_r) + \frac{1}{r}\frac{\partial F_\theta}{\partial\theta} + \frac{\partial F_z}{\partial z}\\ &\quad= \frac{1}{r}\left[r\cos^2\theta\frac{\partial F_x}{\partial x} + r\,\mathrm{sen}\,\theta\cos\theta\left(\frac{\partial F_x}{\partial y} + \frac{\partial F_y}{\partial x}\right) + r\,\mathrm{sen}^2\,\theta\frac{\partial F_y}{\partial y} + \cos\theta F_x + \mathrm{sen}\,\theta F_y\right]\\ &\quad\quad + \frac{1}{r}\left[r\,\mathrm{sen}^2\,\theta\frac{\partial F_x}{\partial x} - r\,\mathrm{sen}\,\theta\cos\theta\left(\frac{\partial F_x}{\partial y} + \frac{\partial F_y}{\partial x}\right) + r\cos^2\theta\frac{\partial F_y}{\partial y}\right.\\ &\quad\quad \left.- \cos\theta F_x - \mathrm{sen}\,\theta F_y\right] + \frac{\partial F_z}{\partial z}\\ &\quad= \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = div F\end{aligned}$$

En el caso de las coordenadas esféricas se puede hacer un procedimiento similar, considerando un punto arbitrario $\mathbf{p} = (r_0, \theta_0, \phi_0)$ en $\mathbb{R}^3$ (en el sistema esférico) y una pequeña región R alrededor de él, tomando incrementos hacia ambos lados de r_0, de θ_0 y de ϕ_0, calculando el flujo del campo $\mathbf{F} = (F_r, F_\theta, F_\phi)$ a través de R, dividiendo luego por el volumen de R y tomando finalmente límite cuando los incrementos mencionados tienden a cero (cuando "R tiende al punto $\mathbf{p}$"). La expresión

resultante es que la divergencia del campo $\mathbf{F} = (F_r, F_\theta, F_\phi) = F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta + F_\phi\mathbf{e}_\phi$ es

$$\operatorname{div}\mathbf{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_r) + \frac{1}{r\operatorname{sen}\phi}\partial\phi(\operatorname{sen}\phi F_\phi) + \frac{1}{r\operatorname{sen}\phi}\frac{\partial F_\theta}{\partial\theta}$$

Por supuesto que también es posible verificar que el segundo miembro de la expresión anterior es en efecto la divergencia del campo **F** usando la correspondencia entre la "versión cartesiana" del campo y su "versión esférica", vista en el apéndice de la sección 2. Es decir, aplicando la expresión del campo **F** en coordenadas esféricas que es

$$\begin{aligned}\mathbf{F}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi) &= \tilde{\mathbf{F}}(r,\theta,\phi)\\ &= (F_r(r,\theta,\phi), F_\theta(r,\theta,\phi), F_\phi(r,\theta,\phi))\\ &= F_r(r,\theta,\phi)\mathbf{e}_r + F_\theta(r,\theta,\phi)\mathbf{e}_\theta + F_\phi(r,\theta,\phi)\mathbf{e}_\phi\end{aligned}$$

donde

$$\begin{aligned}F_r(r,\theta,\phi) &= \operatorname{sen}\phi\cos\theta\tilde{F}_x(r,\theta,\phi) + \operatorname{sen}\theta\operatorname{sen}\phi\tilde{F}_y(r,\theta,\phi) + \cos\phi\tilde{F}_z(r,\theta,\phi)\\ F_\theta(r,\theta,\phi) &= -\operatorname{sen}\theta\tilde{F}_x(r,\theta,\phi) + \cos\theta\tilde{F}_y(r,\theta,\phi)\\ F_\phi(r,\theta.\phi) &= \cos\phi\cos\theta\tilde{F}_x(r,\theta,\phi) + \operatorname{sen}\theta\cos\phi\tilde{F}_y(r,\theta,\phi) - \operatorname{sen}\phi\tilde{F}_z(r,\theta,\phi)\end{aligned}$$

que a su vez

$$\tilde{\mathbf{F}}_{x(y,z)}(r,\theta,\phi) = F_{x(y,z)}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$$

Este es un bonito ejercicio de derivación de funciones compuestas einvitamos al lector a aventurarse en él.

Ejercicios (Capítulo 7, Sección 11)

En los ejercicios 1–4 verifique el teorema de la divergencia con el campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, y el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$

1. $\mathbf{F}(x, y) = (3x + y, 2y)$, $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (3\cos t, 3\operatorname{sen} t)$.
2. $\mathbf{F}(x, y) = (5x - y, x + 4y)$, $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (2\cos t, 3\operatorname{sen} t)$.
3. $\mathbf{F}(x, y) = (x^2, y^2)$, $\boldsymbol{\lambda}$ es un camino cuya imagen es el cuadrado $|x| + |y| = 1$ recorrido positivamente.
4. $\mathbf{F}(x, y) = (2x^2y, xy^3 + y)$, $\boldsymbol{\lambda}$ es un camino cuya imagen es la frontera del triángulo $R = \{(x, y)|0 \le x \le 1, 0 \le y \le 3x\}$ recorrido positivamente.

En los ejercicios 5–10, aplique el teorema de la divergencia para calcular el flujo del campo $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$ a través de la frontera de la región R indicada.

5. $\mathbf{F}(x, y) = (2, 0)$, $R = [0, 3] \times [0, 2]$.
6. $\mathbf{F}(x, y) = (x, y)$, $R = \{(x, y)|0 \le x \le 2, 0 \le y \le x^2\}$.

7. $\mathbf{F}(x, y) = (x^3y, xy^3)$, $R = \{(x, y)|x^2 + y^2 \leq 4\}$.

8. $\mathbf{F}(x, y) = (3x - y^2, x + x^2y)$, $R = \{(x, y)|0 \leq x \leq 2, 0 \leq y \leq x\}$.

9. $\mathbf{F}(x, y) = (3x - y^2, 5x^3 + 2y)$, $R = \{(x, y)|3x^2 + y^2 \leq 4\}$.

10. $\mathbf{F}(x, y) = (x^3 + y^3, 2x^3 + y^3)$, $R = \{(x, y)|1 \leq x^2 + y^2 \leq 4\}$.

Para los campos $\mathbf{F}: \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F}(r, \theta, z) = F_1(r, \theta, z)\mathbf{e}_r + F_2(r, \theta, z)\mathbf{e}_\theta + F_3(r, \theta, z)\mathbf{e}_z$, (en el sistema de coordenadas cilíndricas) en los ejercicios 11–14, determine su divergencia en el punto indicado.

11. $\mathbf{F}(r, \theta, z) = r\mathbf{e}_r + \theta e_\theta + z\mathbf{e}_z$, en el punto $(2, \pi/4, 1)$.

12. $\mathbf{F}(r, \theta, z) = (r\cos\theta)\mathbf{e}_r + \operatorname{sen}\theta\mathbf{e}_\theta + z^2\mathbf{e}_z$, en el punto $(1, 0, 1)$.

13. $\mathbf{F}(r, \theta, z) = z\mathbf{e}_r + r\cos^2\theta\mathbf{e}_\theta + r\theta z\mathbf{e}_z$, en el punto $(3, \pi/2, -1)$.

14. $\mathbf{F}(r, \theta, z) = rz\mathbf{e}_r + (\theta - z^2)\mathbf{e}_\theta + \mathbf{e}_z$, en un punto arbitrario (r, θ, z).

15. $\mathbf{F}(r, \theta, z) = r^3\theta\mathbf{e}_r + rz^4\mathbf{e}_\theta + rz^2\mathbf{e}_z$, en un punto arbitrario (r, θ, z).

Para los campos $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(r, \theta, \phi) = F_1(r, \theta, \phi)\mathbf{e}_r + F_2(r, \theta, \phi)d\theta + F_3(r, \theta, \phi)\, d\phi$, (en el sistema de coordenadas esféricas) en los ejercicios 16–20, determine su divergencia en el punto indicado.

16. $\mathbf{F}(r, \theta, \phi) = \mathbf{e}_r + \mathbf{e}_\theta + \mathbf{e}_\phi$, en el punto $(2, \pi/4, \pi/4)$.

17. $\mathbf{F}(r, \theta, \phi) = r\mathbf{e}_r + \theta\mathbf{e}_\theta + \phi\mathbf{e}_\phi$, en el punto $(1, 3\pi/4, \pi/4)$.

18. $\mathbf{F}(r, \theta, \phi) = r\cos\theta\mathbf{e}_r + r\operatorname{sen}\theta\mathbf{e}_\theta + \operatorname{sen}\phi\mathbf{e}_\phi$, en el punto $(1, \pi/3, \pi/4)$.

19. $\mathbf{F}(r, \theta, \phi) = r\cos\theta\cos\phi\mathbf{e}_r + r\operatorname{sen}\theta\operatorname{sen}\phi\mathbf{e}_\theta + r\cos\phi\mathbf{e}_\phi$, en un punto arbitrario (r, θ, ϕ).

20. $\mathbf{F}(r, \theta, \phi) = (r^2 + \cos\phi)\mathbf{e}_r + (r - \operatorname{sen}\theta)\mathbf{e}_\theta + \operatorname{sen}\phi\mathbf{e}_\phi$, en un punto arbitrario (r, θ, ϕ).

Capítulo 8

Superficies en $\mathbb{R}^3$

En este capítulo abordamos el estudio de ciertos objetos matemáticos en el espacio $\mathbb{R}^3$ conocidos como superficies. Estos objetos matemáticos ya han aparecido anteriormente en el libro (en el capítulo 2), por ejemplo, como gráficas de algunas funciones de 2 variables y como niveles constantes de algunas funciones de 3 variables. Uno de los objetivos de este capítulo será aclarar cuáles gráficas y cuáles funciones de los ejemplos anteriores producen superficies en $\mathbb{R}^3$. El estudio de las superficies en $\mathbb{R}^3$ es por sí mismo interesante y de fundamental importancia para el análisis que la geometría diferencial hace de ellas. No pretendemos aquí incursionar en estos análisis. Lo que haremos en este capítulo es solamente dar la definición de superficie, estudiar algunos casos particulares de ellas y precisar la propiedad de orientabilidad que tienen algunas superficies en $\mathbb{R}^3$. En este sentido el presente es un capítulo menos fuerte que el capítulo 5, donde sí nos aventuramos con algunas profundidades de las curvas en el espacio (curvatura, torsión, fórmulas de Frenet, etc.), las cuales, junto con las superficies, constituyen los elementos de trabajo principales de la geometría diferencial. La razón de esto la descubrirá el lector después de estudiar la definición y ver los primeros resultados que aparecerán aquí: las superficies son, en general, objetos matemáticos más complicados que las curvas. En un curso de cálculo se precisa de ellas pues son las regiones de integración de las llamadas "integrales de superficie" que estudiaremos en el capítulo 9 (al igual que las curvas del capítulo 5 fueron las regiones de integración de las integrales de línea que estudiamos en el capítulo anterior). Lo que presentamos pues en este capítulo es solamente el material necesario para trabajar con las regiones de integración de las integrales del capítulo 9, cuyo estudio nos permitirá cristalizar uno de los objetivos del curso de cálculo en $\mathbb{R}^n$: el teorema de Stokes. Sin embargo, en la última sección de este capítulo presentamos un pequeño estudio de ciertas superficies llamadas *superficies tubulares*, que tienen algunas características interesantes y que son casos particulares de unos objetos matemáticos llamados "tubos", los cuales han sido en los últimos años objetos de investigaciones avanzadas en matemáticas. Esta sección, que presentamos de estudio opcional, la dejamos para los lectores "más matemáticos", esperando dejar sembradas inquietudes para estudios posteriores.

8.1 Superficies simples

En esta sección se tratará de dejar bien establecida la definición de los objetos matemáticos con los que trabajaremos en éste y el próximo capítulos. Estos son las superficies. Ciertamente la palabra "superficie" ha aparecido ya en muchas ocasiones en el texto; sin embargo, ésta ha sido empleada siempre de manera informal, apelando siempre al contenido intuitivo de la palabra. Una situación análoga se presentó cuando, en el capítulo 5, se introdujo el concepto de "curva" (objetos

matemáticos con los que se trabajó en el capítulo 7), aunque esta palabra había sido ya usada desde los primeros cursos de matemáticas elementales. De esta manera, así como en el capítulo 5 se presentó formalmente lo que en cálculo se debe entender por "curva" o "línea", que usamos como regiones de integración en las integrales de línea estudiadas en el capítulo 7, ahora presentamos formalmente el concepto de "superficie", pues queremos estudiar, en el capítulo 9, cierto tipo de integrales (llamadas integrales de superficie), cuyas regiones de integración serán justamente estos objetos matemáticos.

Debemos advertir que hay muchas maneras de abordar el tema de superficies en matemáticas: dependiendo de los intereses posteriores que se tengan en el estudio de ellas, a veces conviene concebirlas de una forma o de otra. La definición que aquí presentamos, que puede ser distinta a la que se presenta en los libros de geometría diferencial, es la que más nos conviene para su posterior usoen el estudio de las integrales de superficie.

De manera informal, una superficie (en $\mathbb{R}^3$) se puede pensar como un objeto "bidimensional" que vive en el espacio $\mathbb{R}^3$, de la misma manera como una curva (en $\mathbb{R}^2$ o $\mathbb{R}^3$) se piensa como un objeto "unidimensional" que vive en el espacio $\mathbb{R}^2$ o $\mathbb{R}^3$. La "unidimensionalidad" de una curva se debe al dominio de la función $\mathbf{f}$ que la define: una curva es la imagen (en $\mathbb{R}^2$ o $\mathbb{R}^3$) de una cierta función definida *en un subconjunto* I de la recta (espacio de dimensión 1). Así, podemos pensar que la función $\mathbf{f}$ es la que se encarga de "meter" el subconjunto I de $\mathbb{R}$ en el espacio y dejar trazada, con sus imágenes, la curva correspondiente (figura 1).

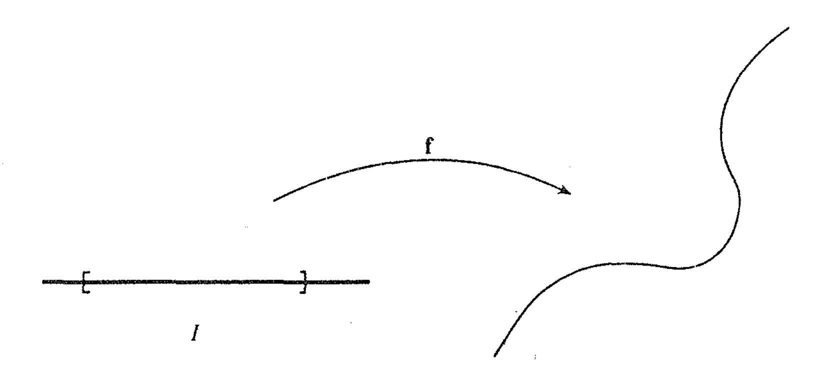


Figura 1. Una curva es un objeto unidimensional en $\mathbb{R}^2$ o $\mathbb{R}^3$)

De la misma manera, una superficie será *la imagen* en $\mathbb{R}^3$ de una cierta función que está definida en un subconjunto S de $\mathbb{R}^2$ (que es bidimensional), y podemos entonces pensar que cada una de estas funciones es una manera de "acomodar" la región S de $\mathbb{R}^2$ en el espacio $\mathbb{R}^3$ (figura 2).

De otro modo: podemos pensar en una superficie como "una sábana en el instante en que ésta va volando por los aires al tender nuestra cama": inicialmente la sábana se encontraba sobre la cama (este es el espacio $\mathbb{R}^2$ y la sábana misma es la región S). Al levantarla de la cama, la sábana cobra vida en el espacio $\mathbb{R}^3$. La forma de ésta en un instante dado es lo que llamaremos "superficie". Por supuesto, algunas acotaciones sobre la región S donde está definida la función $\mathbf{f}$, así como a las características de diferenciabilidad de esta función, deben ser hechas para que podamos aceptar como definición rigurosa esta idea general de superficie. Con nuestro ejemplo de la sábana, debemos imponer condiciones para que al estar volando ésta no aparezcan picos ni aristas en ella.

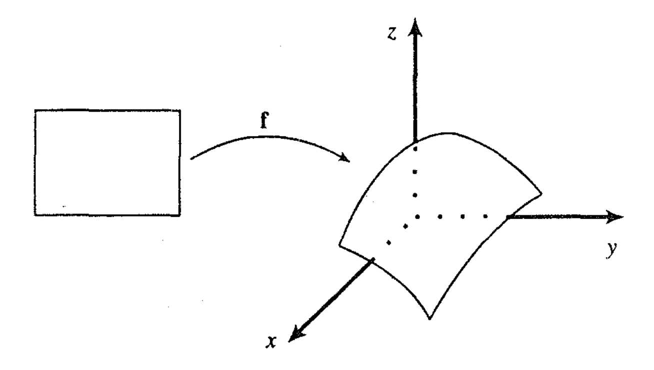


Figura 2. Una superficie es un objeto bidimensional en $\mathbb{R}^3$.

Recordemos que una región $S \subseteq \mathbb{R}^2$ se dice ser del tipo I y del tipo II si ésta se puede describir como:

$$\{(u, v)|a \leq u \leq b, \phi_1(u) \leq v \leq \phi_2(u)\}$$
$$\{(u, v)|c \leq v \leq d, \psi_1(v) \leq u \leq \psi_2(v)\}$$

para algunas funciones continuas $\phi_1, \phi_2: [a, b] \to \mathbb{R}$ y $\psi_1, \psi_2: [c, d] \to \mathbb{R}$ (ver capítulo 6, sección 3).

Definición. Sea $S \subseteq \mathbb{R}^2$ una región del tipo I y del tipo II en $\mathbb{R}^2$ y sea $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (f_1(u, v), f_2(u, v), f_3(u, v))$, una función inyectiva de clase $\mathscr{C}^1$ (es decir, sus funciones coordenadas $f_1, f_2, f_3: S \subset \mathbb{R}^2 \to \mathbb{R}$ son de clase $\mathscr{C}^1$), de modo que los vectores

$$\frac{\partial \mathbf{f}}{\partial u} = \left(\frac{\partial f_1}{\partial u}, \frac{\partial f_2}{\partial u}, \frac{\partial f_3}{\partial u}\right), \quad \frac{\partial \mathbf{f}}{\partial v} = \left(\frac{\partial f_1}{\partial v}, \frac{\partial f_2}{\partial v}, \frac{\partial f_3}{\partial v}\right)$$

son linealmente independientes en todo (u, v) de S. A la imagen de la función $\mathbf{f}$, $K = \mathbf{f}(S)$, se le llama *superficie simple*. ■

El papel que desempeña la propiedad de inyectividad requerida para la función $\mathbf{f}$ es claro: no queremos que las superficies se *autointersecten*, es decir, no queremos que haya dos puntos distintos $\mathbf{p}_1, \mathbf{p}_2 \in S$, de modo que $\mathbf{f}(\mathbf{p}_1) = \mathbf{f}(\mathbf{p}_2) \in K$. (Nuestro ejemplo de la sábana libra automáticamente esta condición, ¿o no?).

Nótese también que la región S, dominio de la función $\mathbf{f}$ cuya imagen ha sido definida como "superficie", es una región *cerrada* y *acotada* del plano $\mathbb{R}^2$ (es decir, es una región compacta de $\mathbb{R}^2$), pues incluye a la frontera que la delimita. Al respecto podríamos objetar sobre lo que se debe entender como que la función $\mathbf{f}$ sea de clase $\mathscr{C}^1$ en los puntos de la frontera de S (todo el estudio sobre diferenciabilidad de las funciones desde el capítulo 2 se ha hecho estando éstas definidas en conjuntos abiertos).[1] La manera de salvar esta situación es pensar en otra función $\tilde{\mathbf{f}}$ que esté definida en un abierto "un poco más grande que S", y que coincida con $\mathbf{f}$ en los puntos de S. A una función $\tilde{\mathbf{f}}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^3$, donde $U \subseteq \mathbb{R}^2$ es un abierto que contiene a S, tal que $\tilde{\mathbf{f}}|_S = \mathbf{f}$, se le llama *extensión* de $\mathbf{f}$ al abierto U. En tal caso los puntos de la frontera de S pasan a ser puntos del abierto U, donde ya

[1] Ver situación análoga en la nota insertada en la primera definición de la sección 2 del capítulo 5

tiene sentido hablar de diferenciabilidad (o bien, de la propiedad de ser de clase $\mathscr{C}^1$) de la función $\mathbf{f}$. Por supuesto, hay muchas maneras de extender a la función $\mathbf{f}: S \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ a un abierto $U \supset S$; sin embargo, es posible demostrar que existe una única manera de efectuar esta extensión de modo que las propiedades diferenciables de $\mathbf{f}$ (la propiedad de ser de clase $\mathscr{C}^1$) se conserve (en todos los puntos de la región S). De este modo, al decir que $\mathbf{f}: S \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ es de clase $\mathscr{C}^1$ en un punto $\mathbf{p}$ de la frontera de S nos estaremos refiriendo a que la extensión mencionada $\tilde{\mathbf{f}}: U \supset S \to \mathbb{R}^3$ es de clase $\mathscr{C}^1$ en el punto $\mathbf{p}$ del abierto U.

Una tercera observación importante sobre la definición de superficie, es acerca de la condición impuesta sobre los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ en los puntos (u, v) de S. Se establece que estos deben ser linealmente independientes. Así como en el capítulo 5 las curvas que llamaron nuestra atención fueron las curvas *regulares*, en las que la derivada (el vector derivada) nunca era nula (siendo esta condición la que garantizaba la no existencia de "picos" o "vértices" en la curva), así mismo, la condición impuesta a los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ en la definición de superficie, tiene la misión (en este caso no tan clara de primera instancia) de no permitir que en la superficie $\mathbf{f}(S) = K$ haya vértices o aristas: las superficies que nos interesan para verlas como regiones de integración en el capítulo siguiente, deben ser *suaves*. Este hecho quedará más claro después de ver algunos ejemplos concretos de superficies que violen esta condición.

Por último diremos que si $\mathbf{f}(S) = K$ es una superficie simple, a la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ (o bien, a sus tres funciones coordenadas $f_1, f_2, f_3: S \subset \mathbb{R}^2 \to \mathbb{R}$), se le llama una *parametrización* de la superficie K.

NOTA: Usaremos las letras (u, v) para denotar a las coordenadas del dominio de la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ y denotaremos como $x = f_1(u, v)$, $y = f_2(u, v)$, $z = f_3(u, v)$ a las funciones coordenadas.

Veamos algunos ejemplos.

Ejemplo 1. Sea S la región $S = \{(u, v)|a \leq u \leq b, c \leq v \leq d\}$ y sea $\mathbf{f}: S \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$. Ciertamente $\mathbf{f}$ es una función inyectiva, de clase $\mathscr{C}^1$ y los vectores

$$\frac{\partial \mathbf{f}}{\partial u} = (1, 0, \alpha) \quad \text{y} \quad \frac{\partial \mathbf{f}}{\partial v} = (0, 1, \beta)$$

son siempre linealmente independientes. Esta es entonces una superficie simple. Obsérvese que la imagen $\mathbf{f}(S) = K$ es (una porción de) un plano que tiene vector normal $(\alpha, \beta, -1)$

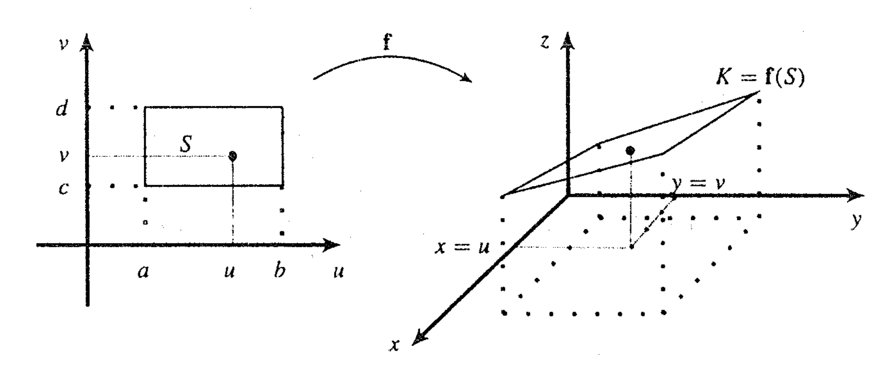


Figura 3. La superficie del ejemplo 1. ■

Ejemplo 2. Sea S la región $S = \{(u, v)|u^2 + v^2 \leq 1\} \subset \mathbb{R}^2$ y sea $\mathbf{f}: S \to \mathbb{R}^3$ la función $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$. Nótese de nuevo que $\mathbf{f}$ es una función inyectiva, de clase $\mathscr{C}^1$ y que los vectores $\frac{\partial \mathbf{f}}{\partial u} = (1, 0, 2u)$, $\frac{\partial \mathbf{f}}{\partial v} = (0, 1, 2v)$ son siempre linealmente independientes. Esta es entonces una superficie simple, la cual representa geométricamente la fracción del paraboloide $z = x^2 + y^2$ que se encuentra por debajo del plano $z = 1$.

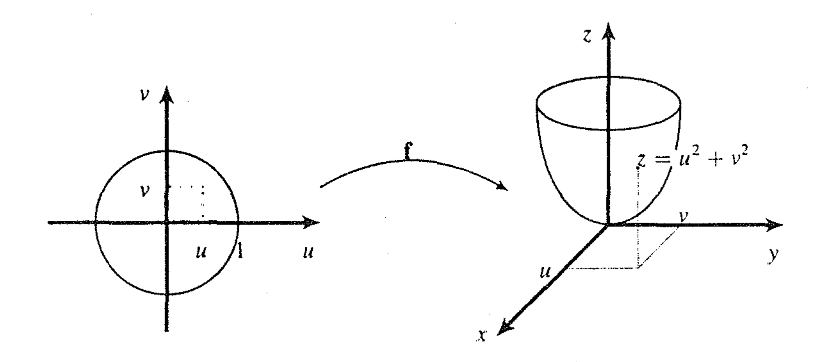


Figura 4. La superficie del ejemplo 2. ■

Los ejemplos anteriores son casos particulares de la siguiente situación más general. Sea $g: S \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ definida en la región S de $\mathbb{R}^2$, la cual es del tipo I y II como las que hemos venido considerando. Hemos visto desde el capítulo 2 que la gráfica de esta función, es decir el conjunto

$$\text{gráfica de } g = \{(x, y, z)|(x, y) \in S, z = g(x, y)\} \subset \mathbb{R}^3$$

es una "superficie" que vive en el espacio $\mathbb{R}^3$. Esta es, en realidad, una superficie simple según la definición que hemos dado en esta sección. En efecto, consideremos la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ dada por $\mathbf{f}(u, v) = (u, v, g(u, v))$. Ciertamente esta es una función inyectiva (¿por qué?), es de clase $\mathscr{C}^1$ (pues sus funciones coordenadas $x = u$, $y = v$, $z = g(u, v)$ lo son) y los vectores

$$\frac{\partial \mathbf{f}}{\partial u} = \left(1, 0, \frac{\partial g}{\partial u}\right), \quad \frac{\partial \mathbf{f}}{\partial v} = \left(0, 1, \frac{\partial g}{\partial v}\right)$$

son siempre linealmente independientes.

Geométricamente esto se ve como la figura 5.

Así pues, las gráficas de funciones $g: S \subset \mathbb{R}^2 \to \mathbb{R}$ de clase $\mathscr{C}^1$ son *un tipo* de superficies simples. Estas, sin embargo, no son "todas" las superficies que nos interesa considerar. Ya desde el capítulo 2 se había considerado que un tipo más general de "superficies" en $\mathbb{R}^3$ lo constituían los niveles constantes de ciertas funciones de 3 variables $h: \mathbb{R}^3 \to \mathbb{R}$. Veamos cuál es la interrelación entre estas superficies de nivel y el concepto de superficie simple que estamos considerando en esta sección.

Sea entonces $h: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ definida en el abierto U de $\mathbb{R}^3$ tal que $h'(\mathbf{p}) \neq 0\ \forall \mathbf{p} \in U$ (es decir, todos los puntos $\mathbf{p} \in U$ son puntos regulares de h; más explícitamente: las tres derivadas parciales de h no se anulan simultáneamente en ningún punto $\mathbf{p} \in U$; de otro modo aún: la función h no tiene puntos críticos). Sea $K = \{(x, y, z) \in U|h(x, y, z) = c\} \subset \mathbb{R}^3$. Es decir, K es el subconjunto de $\mathbb{R}^3$ (más precisamente de U) en el que la función h es constante (e igual a c).

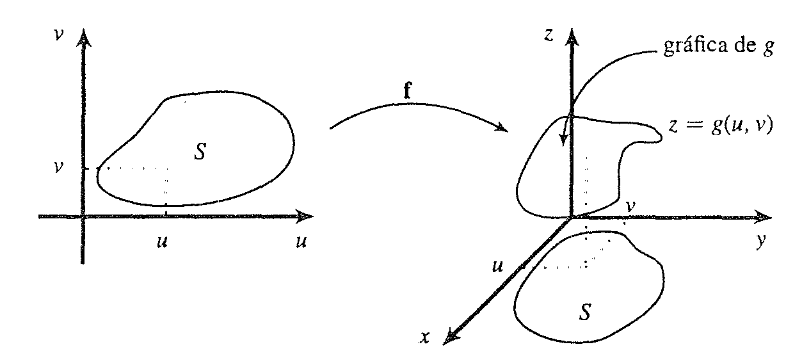


Figura 5. La gráfica de la función g como una superficie en $\mathbb{R}^3$

Supongamos, para fijar ideas, que $\frac{\partial h}{\partial z}(\mathbf{q}) \neq 0$, $\mathbf{q} = (x_0, y_0, z_0) \in U$. Según el teorema de la función implícita (Teorema 3.4.2), es posible, de la expresión $F(x, y, z) = h(x, y, z) - c = 0$, despejar z en términos de x y y, dejando establecida así una función (de clase $\mathscr{C}^1$) $z = \varphi(x, y)$ (de modo que $z_0 = \varphi(x_0, y_0)$) definida en alguna vecindad V del punto (x_0, y_0). De este modo se ve que en los alrededores del punto $\mathbf{q}$, el nivel c de h es la gráfica de una función de clase $\mathscr{C}^1$, $z = \varphi(x, y)$, y es, por lo tanto, una superficie simple. Se tiene pues, que *localmente*, el nivel constante de una función $h\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ de clase $\mathscr{C}^1$ (que tiene solamente puntos regulares) es una superficie simple, gráfica de una función de clase $\mathscr{C}^1$ del tipo $z = \varphi(x, y)$, $x = \psi(y, z)$ o $y = \rho(x, z)$ (dependiendo de si $\frac{\partial h}{\partial z}(\mathbf{q}) \neq 0$, $\frac{\partial h}{\partial x}(\mathbf{q}) \neq 0$ ó $\frac{\partial h}{\partial y}(\mathbf{q}) \neq 0$, respectivamente).

Ejemplo 3. Considere la función $h\colon \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}$, $h(x, y, z) = x^2 + y^2 + z^2$. Esta es una función de clase $\mathscr{C}^1$ y todos los puntos de su dominio son puntos regulares. El nivel constante $K = \{(x, y, z)|x^2 + y^2 + z^2 = c^2\}$ (una esfera con centro en el origen y radio $c > 0$), puede verse, de manera local, como una superficie simple. Por ejemplo, si $c = 1$, en los alrededores del punto $\mathbf{p}_1 = (0, 0, 1)$ (en él $\frac{\partial h}{\partial z}(\mathbf{p}_1) = 2 \neq 0$), K se puede ver como la gráfica de la función, de clase $\mathscr{C}^1$, $z = \varphi(x, y) = \sqrt{1 - x^2 - y^2}$, mientras que en los alrededores del punto $\mathbf{p}_2 = (0, 1, 0)$ (en el que $\frac{\partial h}{\partial y}(\mathbf{p}_2) = 2 \neq 0$), K se puede ver como la gráfica de la función, de clase $\mathscr{C}^1$, $y = \psi(x, y) = \sqrt{1 - x^2 - z^2}$. Estas gráficas son superficies simples.

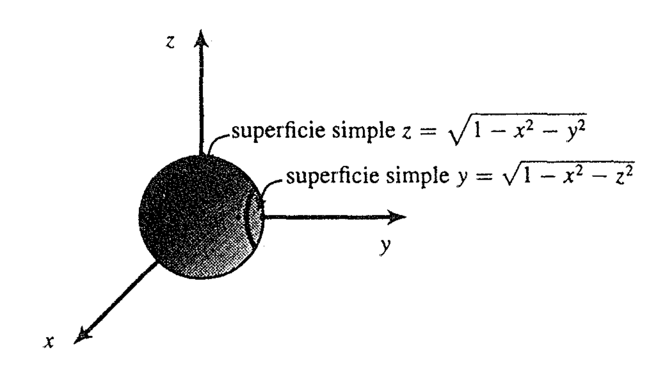


Figura 6. Las superficies del ejemplo 3. ∎

De manera recíproca: una superficie simple K se puede ver localmente como el nivel constante de una función $h\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ de clase $\mathscr{C}^1$ que tiene sólo puntos regulares. En efecto, sea $K = \mathbf{f}(S)$ una superficie simple, imagen de la función inyectiva, de clase $\mathscr{C}^1$, $\mathbf{f}\colon S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f} = (f_1, f_2, f_3)$, tal que los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) = \left(\frac{\partial f_1}{\partial u}(\mathbf{p}), \frac{\partial f_2}{\partial u}(\mathbf{p}), \frac{\partial f_3}{\partial u}(\mathbf{p})\right)$ y $\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = \left(\frac{\partial f_1}{\partial v}(\mathbf{p}), \frac{\partial f_2}{\partial v}(\mathbf{p}), \frac{\partial f_3}{\partial v}(\mathbf{p})\right)$ son linealmente independientes $\forall \mathbf{p} \in S$. Sabemos que estos vectores son linealmente independientes si y sólo si su producto cruz es no nulo, es decir, si y sólo si

$$
\begin{aligned}
(0,0,0) \neq \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial f_1}{\partial u}(\mathbf{p}) & \frac{\partial f_2}{\partial u}(\mathbf{p}) & \frac{\partial f_3}{\partial u}(\mathbf{p}) \\ \frac{\partial f_1}{\partial v}(\mathbf{p}) & \frac{\partial f_2}{\partial v}(\mathbf{p}) & \frac{\partial f_3}{\partial v}(\mathbf{p}) \end{bmatrix} \\
&= \det \begin{bmatrix} \frac{\partial f_2}{\partial u}(\mathbf{p}) & \frac{\partial f_3}{\partial u}(\mathbf{p}) \\ \frac{\partial f_2}{\partial v}(\mathbf{p}) & \frac{\partial f_3}{\partial v}(\mathbf{p}) \end{bmatrix} \mathbf{i} - \det \begin{bmatrix} \frac{\partial f_1}{\partial u}(\mathbf{p}) & \frac{\partial f_3}{\partial u}(\mathbf{p}) \\ \frac{\partial f_1}{\partial v}(\mathbf{p}) & \frac{\partial f_3}{\partial v}(\mathbf{p}) \end{bmatrix} \mathbf{j} + \det \begin{bmatrix} \frac{\partial f_1}{\partial u}(\mathbf{p}) & \frac{\partial f_2}{\partial u}(\mathbf{p}) \\ \frac{\partial f_1}{\partial v}(\mathbf{p}) & \frac{\partial f_2}{\partial v}(\mathbf{p}) \end{bmatrix} \mathbf{k} \\
&= \frac{\partial(f_2, f_3)}{\partial(u,v)}(\mathbf{p})\mathbf{i} - \frac{\partial(f_1, f_3)}{\partial(u,v)}(\mathbf{p})\mathbf{j} + \frac{\partial(f_1, f_2)}{\partial(u,v)}(\mathbf{p})\mathbf{k}
\end{aligned}
$$

Se tiene entonces que, siendo K una superficie simple, los determinantes jacobianos

$$
\frac{\partial(f_i, f_j)}{\partial(u,v)}(\mathbf{p}) \qquad 1 \leq i < j \leq 3
$$

no se pueden anular simultáneamente en punto alguno $\mathbf{p} = (u, v) \in S$. Supongamos, por ejemplo, que $\frac{\partial(f_2, f_3)}{\partial(u,v)}(\mathbf{p}) \neq 0$. Consideremos la función $\mathbf{F}_{2,3}(u, v) = (y, z) = (f_2(u, v), f_3(u, v))$. Esta es una función de clase $\mathscr{C}^1$ con derivada inversible en $\mathbf{p} = (u_0, v_0)$ (es decir, su matriz jacobiana $\mathbf{F}'_{2,3}(\mathbf{p})$ tiene determinante no nulo). Según el Teorema de la Función Inversa (teorema 3.6.1) es posible, de las expresiones $y = f_2(u, v)$, $z = f_3(u, v)$, despejar a u y v en términos de y y z, dejando así establecidas funciones de clase $\mathscr{C}^1$, $u = \eta(y, z)$, $v = \xi(y, z)$ definidas en alguna vecindad V de $(y_0, z_0) = \mathbf{F}_{2,3}(\mathbf{p})$. De este modo en los alrededores del punto (x_0, y_0, z_0), la expresión $x = f_1(u, v) = f_1(\eta(y, z), \xi(y, z))$ es una identidad. Siendo f_1, η y ξ funciones de clase $\mathscr{C}^1$, la función $h(x, y, z) = x - f_1(\eta(y, z), \xi(y, z))$ es de clase $\mathscr{C}^1$ (además, todos los puntos de su dominio son regulares), y su nivel 0 describe localmente, en los alrededores de (x_0, y_0, z_0), la superficie simple K.

Hemos, pues, probado el resultado siguiente.

Teorema 8.1.1

a. Sea $f\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ que tiene en su dominio solamente puntos regulares. Las superficies de nivel de f se pueden ver, localmente, como superficies simples.

b. Sea K una superficie simple. Localmente K está constituida por superficies de nivel de funciones $h\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ de clase $\mathscr{C}^1$ que tienen en su dominio solamente puntos regulares. ∎

Consideremos la superficie simple $K = \mathbf{f}(S)$, parametrizada por la función $\mathbf{f}\colon S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Denotemos por ∂S a la frontera de S y por Int S al interior de S, es decir, Int $S = S - \partial S$. Llamaremos *frontera de la superficie simple* K, denotada por ∂K, al conjunto $\partial K = \mathbf{f}(\partial S)$, y

llamaremos *interior de la superficie simple K*, denotado por Int K, al conjunto Int $K = \mathbf{f}(\text{Int } S)$. Es claro que ∂K e Int K son conjuntos no vacíos (¿por qué?). Más aún, si $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ es una parametrización seccionalmente $\mathscr{C}^1$ de la frontera de la región S, entonces $\mathbf{f} \circ \boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ es una parametrización seccionalmente $\mathscr{C}^1$ de la frontera de K (verifique).

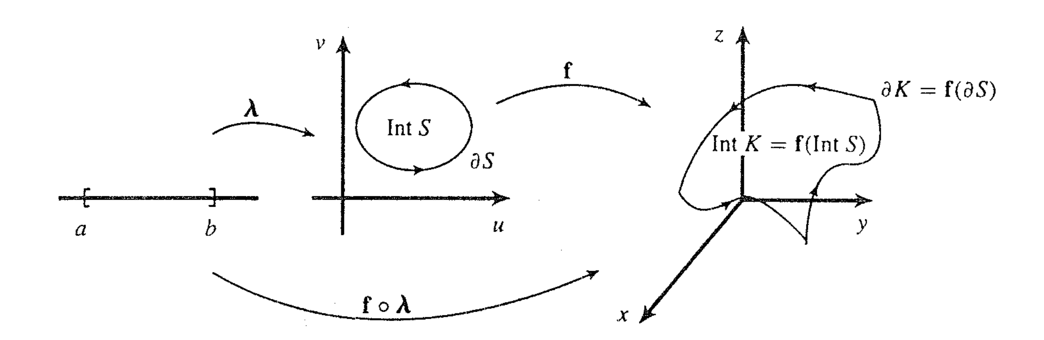


Figura 7. La frontera y el interior de una superficie simple K.

Ejemplo 4. Consideremos la superficie K del ejemplo 1, imagen de la función $\mathbf{f}: S \to \mathbb{R}^3$, $S = \{(u, v) | a \leq u \leq b, c \leq v \leq d\}$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$. La frontera ∂S de S queda parametrizada por el camino $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3 + \boldsymbol{\lambda}_4$, donde

$$\begin{aligned}
&\boldsymbol{\lambda}_1: [a, b] \to \mathbb{R}^2, \quad && \boldsymbol{\lambda}_1(t) = (t, c)\\
&\boldsymbol{\lambda}_2: [c, d] \to \mathbb{R}^2, \quad && \boldsymbol{\lambda}_2(t) = (b, t)\\
&\boldsymbol{\lambda}_3: [a, b] \to \mathbb{R}^2, \quad && \boldsymbol{\lambda}_3(t) = (a + b - t, d)\\
&\boldsymbol{\lambda}_4: [c, d] \to \mathbb{R}^2, \quad && \boldsymbol{\lambda}_4(t) = (a, c + d - t)
\end{aligned}$$

(nótese además que $\boldsymbol{\lambda}$ está orientado positivamente). Entonces la frontera de la superficie simple $K = \mathbf{f}(S)$ es $\partial K = \mathbf{f}(\partial S)$, imagen del camino $\boldsymbol{\mu} = \mathbf{f} \circ \boldsymbol{\lambda} = \mathbf{f} \circ \boldsymbol{\lambda}_1 + \mathbf{f} \circ \boldsymbol{\lambda}_2 + \mathbf{f} \circ \boldsymbol{\lambda}_3 + \mathbf{f} \circ \boldsymbol{\lambda}_4$, donde

$$\begin{aligned}
&\mathbf{f} \circ \boldsymbol{\lambda}_1: [a, b] \to \mathbb{R}^3, \quad && (\mathbf{f} \circ \boldsymbol{\lambda}_1)(t) = \mathbf{f}(t, c) = (t, c, \alpha t + \beta c + \gamma)\\
&\mathbf{f} \circ \boldsymbol{\lambda}_2: [c, d] \to \mathbb{R}^3, \quad && (\mathbf{f} \circ \boldsymbol{\lambda}_2)(t) = \mathbf{f}(b, t) = (b, t, \alpha b + \beta t + \gamma)\\
&\mathbf{f} \circ \boldsymbol{\lambda}_3: [a, b] \to \mathbb{R}^3, \quad && (\mathbf{f} \circ \boldsymbol{\lambda}_3)(t) = \mathbf{f}(a + b - t, d) = (a + b - t, d, \alpha(a + b - t) + \beta d + \gamma)\\
&\mathbf{f} \circ \boldsymbol{\lambda}_4: [c, d] \to \mathbb{R}^3, \quad && (\mathbf{f} \circ \boldsymbol{\lambda}_4)(t) = \mathbf{f}(a, c + d - t) = (a, c + d - t, \alpha a + \beta(c + d - t) + \gamma)
\end{aligned}$$

la cual es, efectivamente, el borde de la porción del plano que representa K. Por otra parte, puesto que el interior de S es el conjunto int $S = \{(u, v) | a < u < b, c < v < d\}$, el interior de la superficie simple K será el conjunto int $K = \{(x, y, z) | z = \alpha x + \beta y + \gamma, a < x < b, c < y < d\}$. ■

Ejemplo 5. Para la superficie simple del ejemplo 2, $K = \mathbf{f}(S)$, donde $\mathbf{f}: S = \{(u, v) | u^2 + v^2 \leq 1\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$, tenemos que $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t)$ es una parametrización de la frontera de S, de modo que $\boldsymbol{\mu} = \mathbf{f} \circ \boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\mu}(t) = \mathbf{f}(\boldsymbol{\lambda}(t)) = \mathbf{f}(\cos t, \operatorname{sen} t) = (\cos t, \operatorname{sen} t, 1)$ es una parametrización de la frontera de K. Observe que la imagen

de μ es el círculo $x^2 + y^2 = 1$, $z = 1$, el cual es efectivamente el borde de la porción del paraboloide $z = x^2 + y^2$ que queda por debajo del plano $z = 1$, que representa K. Por otra parte, el interior de K es el conjunto $\text{Int } K = \mathbf{f}(\text{Int } S) = \mathbf{f}(\{(u, v)|u^2 + v^2 < 1\}) = \{(x, y, z)|z = x^2 + y^2, x^2 + y^2 < 1\}$. ■

Veamos ahora otros ejemplos importantes de superficies simples.

Ejemplo 6. (La semiesfera como superficie simple) Una esfera $x^2 + y^2 + z^2 = 1$ no puede ser una superficie simple, ya que ésta "no tiene frontera" (¿por qué?). Consideremos la semiesfera superior $x^2 + y^2 + z^2 = 1$, $z \geq 0$. El objetivo de este ejercicio es convencernos de que ésta es una superficie simple. Para esto, debemos conseguir una función inyectiva $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$ tal que $\mathbf{f}(S) = K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, z \geq 0\}$. Un primer intento para obtener esta función es pensar en la semiesfera K como la gráfica de la función $f(x, y) = \sqrt{1 - x^2 - y^2}$, definida en la región $S = \{(x, y)|x^2 + y^2 \leq 1\}$. Observe, sin embargo, que en los puntos de la frontera de S, es decir, en los puntos (x, y) tales que $x^2 + y^2 = 1$, la función f no tiene derivadas parciales y por lo tanto esta función no es de clase $\mathscr{C}^1$ en S. Otro intento para obtener la función $\mathbf{f}$ es pensando la semiesfera K descrita en coordenadas esféricas, por medio de la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, definida en la región $S = \{(u, v)|0 \leq u \leq 2\pi, 0 \leq v \leq \pi/2\}$ (que geométricamente representa un rectángulo en el plano uv -las letras u y v que estamos usando corresponden a las tradicionales letras θ y φ, respectivamente, que aparecen en las coordenadas esféricas), dada por

$$\mathbf{f}(u, v) = (\cos u \operatorname{sen} v, \operatorname{sen} u \operatorname{sen} v, \cos v)$$

Esta función es claramente de clase $\mathscr{C}^1$, pero no es inyectiva ($\mathbf{f}(0, v) = (\operatorname{sen} v, 0, \cos v) = \mathbf{f}(2\pi, v)$). Así pues, esta función $\mathbf{f}$ tampoco sirve como parametrización de la semiesfera K.

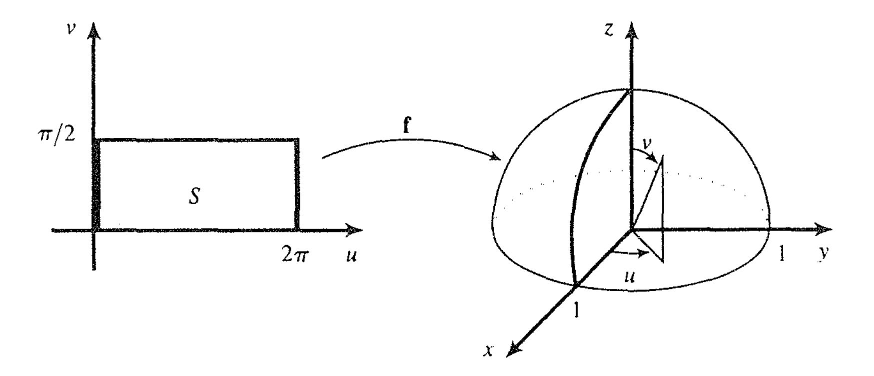


Figura 8. La función $\mathbf{f}(u, v) = (\cos u \cos v, \operatorname{sen} u \operatorname{sen} v, \cos v)$.

La manera de obtener la función $\mathbf{f}$ que parametrice a la semiesfera superior K es por medio de la llamada "proyección estereográfica": defina la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, donde $S = \{(u, v)|u^2+v^2 \leq 1\}$, como $\mathbf{f}(u, v) = (x, y, z) =$ punto de la esfera $x^2 + y^2 + z^2 = 1$ que es colineal con el punto $(0, 0, -1)$ y $(u, v, 0)$. Geométricamente (figura 9)

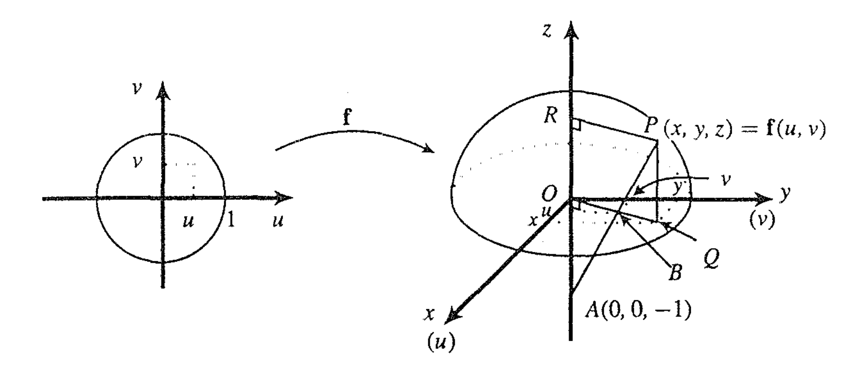


Figura 9. La proyección estereográfica.

Es claro que los puntos O, B y Q son colineales. Entonces se tiene la relación

$$uy = vx \tag{1}$$

Por otra parte, los triángulos ARP y AOB son semejantes, por lo que

$$\frac{|\overline{OB}|}{|\overline{OA}|} = \frac{|\overline{RP}|}{|\overline{RA}|} \tag{2}$$

donde: $|\overline{OB}| = \sqrt{u^2 + v^2}$, $|\overline{OA}| = 1$, $|\overline{RP}| = |\overline{OQ}| = \sqrt{x^2 + y^2}$, $|\overline{RA}| = 1 + z$, quedando entonces (2) escrita como

$$u^2 + v^2 = \frac{x^2 + y^2}{(1 + z)^2}$$

Combinando ésta con (1) se obtienen las expresiones

$$u = \frac{x}{1 + z}, v = \frac{y}{1 + z} \tag{3}$$

en que $z = \sqrt{1 - x^2 - y^2}$. Sustituyendo $x = u(1 + z)$, $y = v(1 + z)$ en $x^2 + y^2 + z^2 = 1$ y resolviendo para z se obtiene que (recordando que $z \geq 0$)

$$z = \frac{1 - u^2 - v^2}{1 + u^2 + v^2}$$

con la cual se llega también (usando (3)) a

$$x = \frac{2u}{1 + u^2 + v^2}, \quad y = \frac{2v}{1 + u^2 + v^2}$$

Entonces la función $\mathbf{f}: S \to \mathbb{R}^3$ es:

$$\mathbf{f}(u, v) = \frac{1}{1 + u^2 + v^2}(2u, 2v, 1 - u^2 - v^2)$$

la cual es inyectiva, de clase $\mathscr{C}^1$ y tal que los vectores

$$\frac{\partial \mathbf{f}}{\partial u} = \frac{1}{(1+u^2+v^2)^2}(2(1-u^2+v^2), -4uv, -4u)$$
$$\frac{\partial \mathbf{f}}{\partial v} = \frac{1}{(1+u^2+v^2)^2}(-4uv, 2(1+u^2-v^2), -4v)$$

son siempre linealmente independientes $\forall (u, v) \in S$. Esta función es entonces una parametrización de la semiesfera superior K. ■

Ejemplo 7. La gráfica de la función $g: \mathbb{R}^2 \to \mathbb{R}$, $g(x, y) = |x| + |y|$ es una superficie en $\mathbb{R}^3$ que no es simple. De hecho, una función $\mathbf{f}: \mathbb{R}^2 \to \mathbb{R}^3$ que tiene a tal gráfica como imagen es $\mathbf{f}(u, v) = (u, v, |u| + |v|)$, la cual claramente carece de propiedades de diferenciabilidad en todos los puntos $(u, v) \in \mathbb{R}^2$ en los que $u = 0$ y/o $v = 0$. Sin embargo, podemos tener funciones $\mathbf{f}: \mathbb{R}^2 \to \mathbb{R}^3$ *de clase* $\mathscr{C}^1$, cuya imagen sea la gráfica de g, por ejemplo $\tilde{\mathbf{f}}(u, v) = (u^3, v^3, u^2|u| + v^2|v|)$. Obsérvese que la tercera función coordenada de $\tilde{\mathbf{f}}$ es la suma de los valores absolutos de cada una de sus dos primeras coordenadas; es decir, se tiene $z = |x| + |y|$ donde $\tilde{\mathbf{f}}(u, v) = (x, y, z)$. Así, la imagen de $\tilde{\mathbf{f}}$ es la misma gráfica de g. Ciertamente las funciones $x = u^3$, $y = v^3$, $z = u^2|u| + v^2|v|$ son de clase $\mathscr{C}^1$ en todo $(u, v) \in \mathbb{R}^2$, pero esto no significa aún que la gráfica de g sea una superficie simple (y que $\tilde{\mathbf{f}}$ sea una parametrización de ella); las derivadas parciales de $\tilde{\mathbf{f}}$ son

$$\frac{\partial \tilde{\mathbf{f}}}{\partial u} = (3u^2, 0, 3u|u|), \qquad \frac{\partial \tilde{\mathbf{f}}}{\partial v} = (0, 3v^2, 3v|v|)$$

Nótese que estos vectores son linealmente independientes siempre que $u \neq 0$ y $v \neq 0$. Si $u = 0$ ó $v = 0$, los vectores son linealmente dependientes (uno de ellos es el vector cero) y por lo tanto la imagen de $\tilde{\mathbf{f}}$ no puede ser una superficie simple según se definió en esta sección. Se tiene entonces que aunque la gráfica de $g: \mathbb{R}^2 \to \mathbb{R}$ sea la imagen de una función $\mathbf{f}: \mathbb{R}^2 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$, esta gráfica puede no ser una superficie simple (situación análoga en curvas: aunque la gráfica de $g: \mathbb{R} \to \mathbb{R}$, $g(x) = |x|$ es la imagen del camino $\boldsymbol{\lambda}: \mathbb{R} \to \mathbb{R}^2$ diferenciable $\boldsymbol{\lambda}(t) = (t^3, t^2|t|)$, esta gráfica no es una curva regular). De modo geométrico, la gráfica de g es un "cono cuadrado" con vértice en el origen, cuyo nivel $c > 0$ corresponde al cuadrado $|x| + |y| = c$

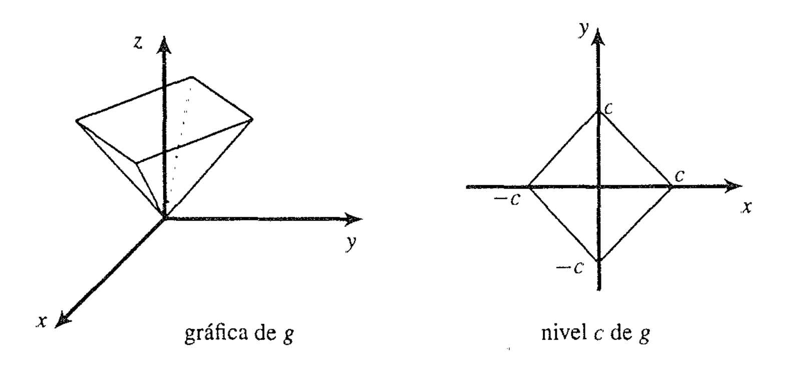


Figura 10. El "cono cuadrado" $z = |x| + |y|$ ■

Diremos que la superficie K es *seccionalmente simple* (o simple en partes) si es la imagen de una función inyectiva $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$ con derivadas parciales $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ linealmente independientes en todo S, donde S es una región de $\mathbb{R}^2$ que se puede descomponer como una unión de subconjuntos S_i, $i = 1, 2, \ldots, n$ que son del tipo I y del tipo II a la vez.

Ejemplo 8. El cilindro $K = \{(x, y, z)|x^2+y^2 = 1, -1 \leq z \leq 1\}$ es una superficie seccionalmente simple, pues es la imagen de la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, definida en $S = \{(u, v)|1 \leq u^2 + v^2 \leq 9\}$

$$\mathbf{f}(u, v) = \left(\frac{u}{\sqrt{u^2+v^2}}, \frac{v}{\sqrt{u^2+v^2}}, \sqrt{u^2+v^2} - 2\right)$$

Esta es una función inyectiva, de clase $\mathscr{C}^1$, y sus parciales

$$\frac{\partial \mathbf{f}}{\partial u} = \left(\frac{v^2}{(u^2+v^2)^{3/2}}, \frac{-uv}{(u^2+v^2)^{3/2}}, \frac{u}{\sqrt{u^2+v^2}}\right)$$

$$\frac{\partial \mathbf{f}}{\partial v} = \left(\frac{-uv}{(u^2+v^2)^{3/2}}, \frac{u^2}{(u^2+v^2)^{3/2}}, \frac{v}{\sqrt{u^2+v^2}}\right)$$

son linealmente independientes para todo $(u, v) \in S$. Obsérvese que el conjunto S donde se define $\mathbf{f}$ es un anillo circular de radios 1 y 3, el cual puede ser descompuesto como unión de subconjuntos que son del tipo I y del tipo II a la vez (dejemos como ejercicio para el lector que llene los detalles de esta afirmación).

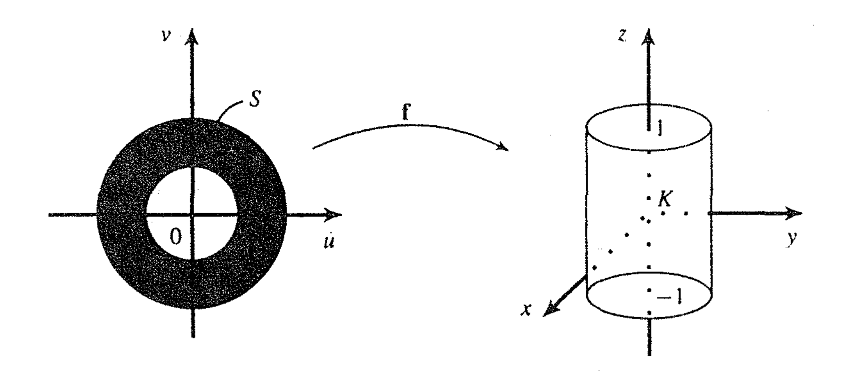


Figura 11. La sección S y la superficie K del ejemplo 8. ■

Ejercicios (Capítulo 8, Sección 1)

En los ejercicios 1–10 se da un conjunto $K \subset \mathbb{R}^3$. Demuestre que K es una superficie simple, dando una función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ cuya imagen sea K.

a. Verifique que $\mathbf{f}$ es una parametrización de K, mostrando que $\mathbf{f}$ es una función inyectiva, de clase $\mathscr{C}^1$, y que sus vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ son linealmente independientes para todo $(u, v) \in S$.

b. Describa explícitamente la frontera de K, dando una parametrización $\mu: [a, b] \to \mathbb{R}^3$ de ella.

c. Describa explícitamente el interior de K.

d. Haga un bosquejo de K.

1. $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, 1/2 \leq z \leq 1\}$.

2. $K = \{(x, y, z)|2x^2 + y^2 + z^2 = 2, 1/2 \leq x \leq 1\}$.

3. $K = \{(x, y, z)|x^2 + 3y^2 + 3z^2 = 3, -1 \leq y \leq -1/3\}$.

4. $K = \{(x, y, z)|z = 3x^2 + y^2, 3x^2 + y^2 \leq 4\}$.

5. $K = \{(x, y, z)|z = x^2 + y^2, x^2 + y^2 \leq 4x\}$.

6. $K = \{(x, y, z)|z = e^{-(x^2+y^2)}, x^2 + y^2 \leq 1\}$.

7. $K = \{(x, y, z)|z = x^2 + y^2 + 3x - 8y + 1, x^2 + y^2 \leq 3y\}$.

8. $K = \{(x, y, z)|z = -x^2 - y^2, 0 \leq x \leq 1, 0 \leq y \leq x\}$.

9. $K = \{(x, y, z)|x - y^2 - z^2 = 1, y^2 + z^2 \leq 4\}$.

10. $K = \{(x, y, z)|y = x^2 + z^2 + 2x + 4z + 8, 0 \leq x \leq 1, 0 \leq z \leq 1 - x\}$.

11. Compruebe que los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$, derivadas parciales de la función que parametriza la semiesfera superior $z = \sqrt{1 - x^2 - y^2}$ (con la proyección estereográfica, ver ejemplo 6), son linealmente independientes, calculando su producto cruz (y verificando que es no nulo para todo $(u, v) \in S = \{(u, v)|u^2 + v^2 \leq 1\}$).

12. Considere la superficie simple $K = \mathbf{f}(S)$, parametrizada por la función $\mathbf{f}: S = [0, \pi] \times [-1, 1] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (\cos u, \operatorname{sen} u, v)$.

a. Compruebe que la función $\mathbf{f}$ satisface todas las propiedades requeridas en la definición de superficie simple.

b. Describa geométricamente K.

c. Describa la frontera (dando una parametrización de ella) y el interior de K.

13. Considere la superficie simple $K = \mathbf{g}(S)$, parametrizada por la función $\mathbf{g}: S = [-\pi/2, \pi/2] \times [-1, 1] \to \mathbb{R}^3$, $\mathbf{g}(u, v) = (-\operatorname{sen} u, \cos u, v)$.

a. Compruebe que la función $\mathbf{g}$ satisface todas las propiedades requeridas en la definición de superficie simple.

b. Describa K en forma geométrica.

c. Describa la frontera (dando una parametrización de ella) y el interior de K.

d. Compare con el ejercicio anterior.

14. Considere la función $\mathbf{f}$ del ejercicio 12 y la función $\mathbf{g}$ del ejercicio 13. Se ha visto que ambas describen *la misma* superficie simple K. Determine una función biyectiva $\varphi: [-\pi/2, \pi/2] \times [-1, 1] \to [0, \pi] \times [-1, 1]$ tal que $\mathbf{g} = \mathbf{f} \circ \varphi$. Compruebe que la derivada de φ (es decir, la matriz jacobiana que representa a la derivada de φ) tiene determinante no nulo.

15. Sean K_1 y K_2 dos superficies simples. ¿Es la unión $K_1 \cup K_2$ una superficie simple? Si lo es, demuéstrelo; caso contrario, dé un contraejemplo.

16. Sean K_1 y K_2 dos superficies simples. ¿Es la intersección $K_1 \cap K_2$ una superficie simple? Si lo es, demuéstrelo; caso contrario, dé un contraejemplo.

17. Demuestre que $K = \{(x, y, z)|z = x^2 + y^2, 1 \leq x^2 + y^2 \leq 4\}$ es una superficie seccionalmente simple. Describa K en forma geométrica.

8.2 Reparametrizaciones

Así como una curva en $\mathbb{R}^2$ o en $\mathbb{R}^3$ puede ser la imagen de distintas funciones $\boldsymbol{\alpha}: I \subseteq \mathbb{R} \to \mathbb{R}^2$ (o $\mathbb{R}^3$), una superficie K en $\mathbb{R}^3$ también puede ser descrita como imagen de diferentes funciones $\mathbf{f}$ que la parametrizan.

Ejemplo 1. Consideremos la superficie K del ejemplo 1 de la sección anterior, imagen de la función $\mathbf{f}: [a, b] \times [c, d] \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$. Sea $\varphi: [a', b'] \times [c', d'] \subset \mathbb{R}^2 \to \mathbb{R}^2$ la función

$$\varphi(s, t) = \left(\frac{b-a}{b'-a'}(s-a') + a, \frac{d-c}{d'-c'}(t-c') + c\right)$$

Nótese que esta función transforma el rectángulo $[a', b'] \times [c', d']$ en el rectángulo $[a, b] \times [c, d]$ de manera biyectiva

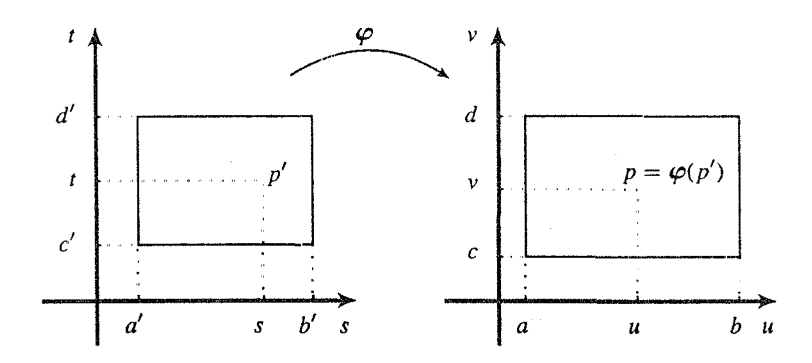


Figura 1. La función φ, que convierte $[a', b'] \times [c', d']$ en $[a, b] \times [c, d]$

La función $\mathbf{g} = \mathbf{f} \circ \varphi: [a', b'] \times [c', d'] \to \mathbb{R}^3$

$$\begin{aligned}
\mathbf{g}(s, t) &= (\mathbf{f} \circ \varphi)(s, t) = \mathbf{f}(\varphi(s, t)) \\
&= \mathbf{f}\left(\frac{b-a}{b'-a'}(s-a') + a, \frac{d-c}{d'-c'}(t-c') + c\right) \\
&= \left(\frac{b-a}{b'-a'}(s-a') + a, \frac{d-c}{d'-c'}(t-c') + c,\right. \\
&\quad \left.\alpha\left(\frac{b-a}{b'-a'}(s-a') + a\right) + \beta\left(\frac{d-c}{d'-c'}(t-c') + c\right) + \gamma\right)
\end{aligned}$$

tiene la misma imagen que **f**, es inyectiva, de clase $\mathscr{C}^1$ y sus derivadas parciales

$$\frac{\partial \mathbf{g}}{\partial s} = \left(\frac{b-a}{b'-a'}, 0, \alpha\left(\frac{b-a}{b'-a'}\right)\right), \quad \frac{\partial \mathbf{g}}{\partial t} = \left(0, \frac{d-c}{d'-c'}, \beta\left(\frac{d-c}{d'-c'}\right)\right)$$

son siempre linealmente independientes. Se trata entonces de una función que parametriza a la superficie (al plano) $K = \mathbf{f}([a,b] \times [c,d]) = \mathbf{g}([a',b'] \times [c',d'])$. Esquemáticamente

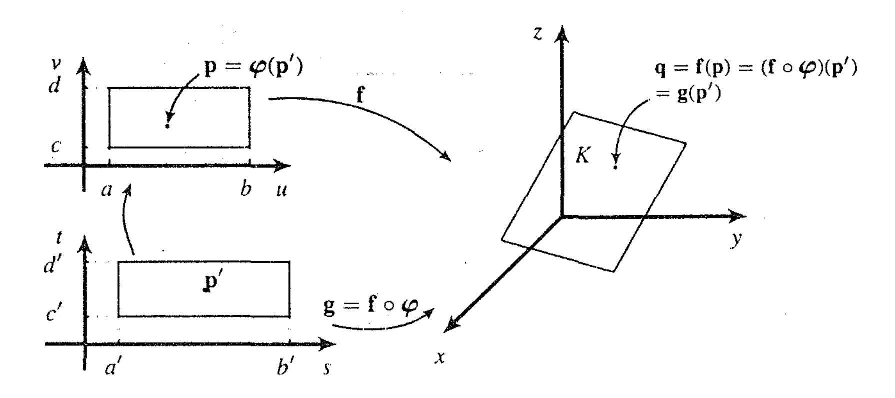


Figura 2. Las funciones **f**, φ, y $\mathbf{g} = \mathbf{f} \circ \varphi$ ■

A una función **g** como la del ejemplo anterior, le llamaremos *reparametrización* de la función **f** (también diremos que es una reparametrización de la superficie K). Obsérvese que las características de la función φ que permitieron que la imagen de la función $\mathbf{g} = \mathbf{f} \circ \varphi$ siguiera siendo una superficie simple fueron: 1) φ mandó de manera biyectiva la región $S' = \{(s,t)|a' \leq s \leq b', c' \leq t \leq d'\}$ en la región $S = \{(u,v)|a \leq u \leq b, c \leq v \leq d\}$, con lo cual conservó, por una parte, la inyectividad de **f**, y por otra, se aseguró que $\mathbf{f}(S) = \mathbf{g}(S')$; 2) La función φ es de clase $\mathscr{C}^1$, con lo cual se asegura que la función $\mathbf{g} = \mathbf{f} \circ \varphi$, siendo **f** de clase $\mathscr{C}^1$, sea también de clase $\mathscr{C}^1$; 3) La función φ tiene una derivada $\varphi'(s,t)$ (la cual es una matriz 2×2) que permite conservar la propiedad de independencia lineal de los vectores $\frac{\partial \mathbf{f}}{\partial u}$, $\frac{\partial \mathbf{f}}{\partial v}$ al momento de hacer la composición con **f**, de modo que los vectores $\frac{\partial \mathbf{g}}{\partial s} = \frac{\partial}{\partial s}(\mathbf{f} \circ \varphi)$ y $\frac{\partial \mathbf{g}}{\partial t} = \frac{\partial}{\partial t}(\mathbf{f} \circ \varphi)$ siguen siendo linealmente independientes en todos los puntos $(s,t) \in S'$.

De este modo, las propiedades de la función **f** que describe a la superficie $K = \mathbf{f}(S)$, a saber, inyectividad, ser de clase $\mathscr{C}^1$ y tener los vectores de sus derivadas parciales linealmente independientes, se conservan al momento de hacer la composición $\mathbf{g} = \mathbf{f} \circ \varphi$.

Definición. Sea K una superficie simple, imagen de la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, y sea $\varphi: S' \subset \mathbb{R}^2 \to S \subset \mathbb{R}^2$ una función biyectiva definida en la región $S' \subset \mathbb{R}^2$ la cual es del tipo I y II a la vez, de clase $\mathscr{C}^1$ de modo que su derivada $\varphi'(s,t)$ es siempre inversible, es decir,

$$\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)} = \det \begin{bmatrix} \dfrac{\partial \varphi_1}{\partial s} & \dfrac{\partial \varphi_1}{\partial t} \\ \dfrac{\partial \varphi_2}{\partial s} & \dfrac{\partial \varphi_2}{\partial t} \end{bmatrix} \neq 0 \qquad \forall (s,t) \in S'$$

donde $\varphi_1, \varphi_2: S' \subset \mathbb{R}^2 \to \mathbb{R}$ son las funciones coordenadas de φ. A la función compuesta $\mathbf{g} = \mathbf{f} \circ \varphi: S' \subset \mathbb{R}^2 \to \mathbb{R}^3$ se le llama *reparametrización de* **f** (o de K)

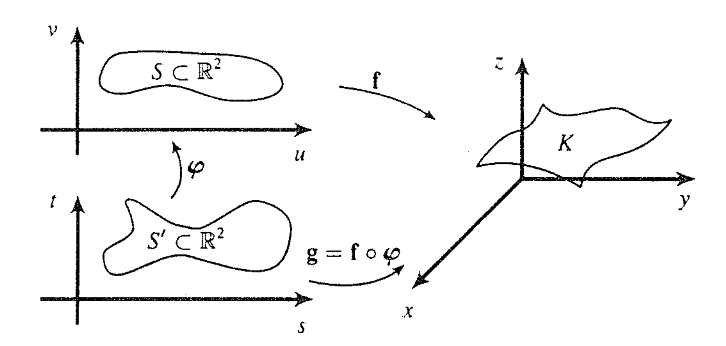


Figura 3. La reparametrización de f (o de K).

Según la definición anterior, una reparametrización $\mathbf{g}$ de $\mathbf{f}$, es de hecho una función que parametriza a la misma superficie K. En efecto, el hecho de que $\mathbf{g}(S') = \mathbf{f}(S) = K$ es claro. Que la función $\mathbf{g}$ es de clase $\mathscr{C}^1$ se deduce de que es una composición de dos funciones de clase $\mathscr{C}^1$. Falta verificar que los vectores $\frac{\partial \mathbf{g}}{\partial s}$ y $\frac{\partial \mathbf{g}}{\partial t}$ son linealmente independientes en todo $(s, t) \in S'$. Se tiene que $\mathbf{g}(s, t) = (\mathbf{f} \circ \varphi)(s, t)$, de modo que, según la regla de la cadena

$$\underset{\text{matriz de } 3\times 2}{\underset{\uparrow}{\mathbf{g}'(s,t)}} = \underset{\text{matriz de } 3\times 2}{\underset{\uparrow}{\mathbf{f}'(\varphi(s,t))}}\ \underset{\text{matriz de } 2\times 2}{\underset{\uparrow}{\varphi'(s,t)}}$$

Esquemáticamente

$$\begin{bmatrix} \uparrow & & \uparrow \\ \frac{\partial \mathbf{g}}{\partial s}(s,t) & \cdots & \frac{\partial \mathbf{g}}{\partial t}(s,t) \\ \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & & \uparrow \\ \frac{\partial \mathbf{f}}{\partial u}(\varphi(s,t)) & \cdots & \frac{\partial \mathbf{f}}{\partial v}(\varphi(s,t)) \\ \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} a(s,t) & b(s,t) \\ c(s,t) & d(s,t) \end{bmatrix}$$

donde $a(s, t) = \frac{\partial \varphi_1}{\partial s}(s, t)$, $b(s, t) = \frac{\partial \varphi_1}{\partial t}(s, t)$, $c(s, t) = \frac{\partial \varphi_2}{\partial s}(s, t)$, $d(s, t) = \frac{\partial \varphi_2}{\partial t}(s, t)$, y $a(s, t)d(s, t) - b(s, t)c(s, t) \neq 0\ \forall (s, t) \in S'$.

Entonces

$$\left.\begin{aligned} \frac{\partial \mathbf{g}}{\partial s}(s,t) &= a(s,t)\frac{\partial \mathbf{f}}{\partial u}(\varphi(s,t)) + c(s,t)\frac{\partial \mathbf{f}}{\partial v}(\varphi(s,t)) \\ \frac{\partial \mathbf{g}}{\partial t}(s,t) &= b(s,t)\frac{\partial \mathbf{f}}{\partial u}(\varphi(s,t)) + d(s,t)\frac{\partial \mathbf{f}}{\partial v}(\varphi(s,t)) \end{aligned}\right\} \tag{A}$$

de modo que

$$\frac{\partial \mathbf{g}}{\partial s}(s,t) \times \frac{\partial \mathbf{g}}{\partial t}(s,t) = (a(s,t)d(s,t) - c(s,t)b(s,t))\frac{\partial \mathbf{f}}{\partial u}(\varphi(s,t)) \times \frac{\partial \mathbf{f}}{\partial v}(\varphi(s,t))$$

Puesto que el escalar que aparece multiplicando al producto cruz del lado derecho de esta expresión es distinto de cero para todo $(s, t) \in S'$ (es de hecho el jacobiano $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)}$) y el producto cruz $\frac{\partial \mathbf{f}}{\partial u}(\varphi(s, t)) \times \frac{\partial \mathbf{f}}{\partial v}(\varphi(s, t))$ es, por hipótesis, no nulo (ya que $\mathbf{f}$ es una función que parametriza a K), concluimos que el producto cruz $\frac{\partial \mathbf{g}}{\partial s}(s, t) \times \frac{\partial \mathbf{g}}{\partial t}(s, t)$ es no nulo para todo $(s, t) \in S'$, por lo que estos vectores $\frac{\partial \mathbf{g}}{\partial s}$ y $\frac{\partial \mathbf{g}}{\partial t}$ son siempre linealmente independientes.

La última fórmula obtenida tiene un contenido geométrico importante sobre el que se discutirá la próxima sección. Queremos, por el momento, dejarla marcada para luego hacer referencia a ella. La escribiremos como

$$\frac{\partial \mathbf{g}}{\partial s}(s,t) \times \frac{\partial \mathbf{g}}{\partial t}(s,t) = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)} \frac{\partial \mathbf{f}}{\partial u}(\varphi(s,t)) \times \frac{\partial \mathbf{f}}{\partial v}(\varphi(s,t)) \qquad (\hat{\text{A}})$$

Ejemplo 2. Cuando se tiene una superficie simple $K = \mathbf{f}(S)$ imagen de una función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ definida en un rectángulo $S = [a, b] \times [c, d]$, es posible decir que, " sin pérdida de generalidad", podemos suponer que la función $\mathbf{f}$ está definida en el cuadrado $S' = [0, 1] \times [0, 1]$. Esto significa que podemos tener una reparametrización $\mathbf{g}: S' \subset \mathbb{R}^2 \to \mathbb{R}^3$ de la superficie K, definida en el cuadrado S'. De hecho, se verifica fácilmente que $\mathbf{g}: [0, 1] \times [0, 1] \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{g}(s, t) = \mathbf{f}((b-a)s+a, (d-c)t+c)$ es la reparametrización mencionada. ■

Ejemplo 3. Consideremos la superficie $K = \mathbf{f}(S)$, en la que $\mathbf{f}: [-1, 1] \times [-1, 1] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, v^2)$. Se trata de un cilindro parabólico de altura 1 y longitud 2, gráfica de $z = y^2$, $-1 \leq y \leq 1, -1 \leq x \leq 1$.

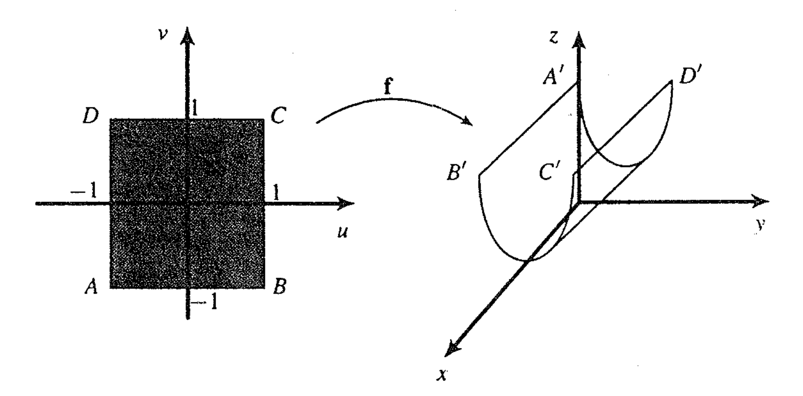


Figura 4. El cilindro parabólico $z = y^2$ en la región $[-1, 1] \times [-1, 1]$.

Nótese que la frontera de K, imagen de la frontera de S, está constituida por las líneas

$$\overline{A'B'} = \mathbf{f}(\overline{AB}) = \{(x, y, z) | -1 \leq x \leq 1, y = -1, z = 1\}$$
$$\overline{C'D'} = \mathbf{f}(\overline{CD}) = \{(x, y, z) | -1 \leq x \leq 1, y = 1, z = 1\}$$
$$\overline{B'C'} = \mathbf{f}(\overline{BC}) = \{(x, y, z) | x = 1, -1 \leq y \leq 1, z = y^2\}$$
$$\overline{A'D'} = \mathbf{f}(\overline{AD}) = \{(x, y, z) | x = -1, -1 \leq y \leq 1, z = y^2\}$$

Sea $\varphi: [-1, 1] \times [-1/k, 1/k] \to [-1, 1] \times [-1, 1]$, $\varphi(s, t) = (kt, -s)$, donde $k \in \mathbb{R}$, $k \neq 0$. Ciertamente φ es una biyección, es de clase $\mathscr{C}^1$, y su derivada $\varphi'(s, t) = \begin{bmatrix} 0 & k \\ -1 & 0 \end{bmatrix}$ es siempre una matriz inversible (se tiene de hecho, $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)} = k \neq 0$).

Entonces $\mathbf{g} = \mathbf{f} \circ \varphi: [-1, 1] \times [-1/k, 1/k] \to \mathbb{R}^3$,

$$\mathbf{g}(s, t) = (\mathbf{f} \circ \varphi)(s, t) = \mathbf{f}(\varphi(s, t)) = \mathbf{f}(kt, -s) = (kt, -s, s^2)$$

es una reparametrización de K. Obsérvese que con esta parametrización, la frontera de K, definida como la imagen de la frontera de $S' = [-1, 1] \times [-1/k, 1/k]$ bajo la función $\mathbf{g}$, es la misma que la considerada anteriormente ($\partial K = \mathbf{f}(\partial S)$, $S = [-1, 1] \times [-1, 1]$). Por ejemplo, los puntos $-1 \leq s \leq 1$, $t = 1/k$ de la frontera de S' van a dar, bajo la parametrización $\mathbf{g}$, a los puntos del arco de parábola $B'C'$. Dejamos al lector que verifique en detalle este hecho. ■

La situación mostrada en el ejemplo anterior acontece en general: la frontera de una superficie es la imagen de la frontera del dominio de *cualquier* parametrización de ella. Esto mismo para el interior de la superficie. Es decir, $\partial K = \mathbf{f}(\partial S)$ e $\operatorname{Int} K = \mathbf{f}(\operatorname{Int} S)$, donde $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ es cualquier parametrización de K. La demostración de este hecho, que requiere de muchos cuidados técnicos y detallados, queda como ejercicio para los lectores que gusten de este tipo de aventuras matemáticas.

Ejercicios (Capítulo 8, Sección 2)

Sea $K = \mathbf{f}(S)$ una superficie simple, parametrizada por la función $\mathbf{f}: S \to \mathbb{R}^3$, definida en la región $S = \{(u, v) | u^2 + v^2 \leq 1\}$. Determine con cuáles de las funciones $\varphi: S' \subset \mathbb{R}^2 \to \mathbb{R}^2$ dadas en los ejercicios 1–7, se pueden obtener reparametrizaciones $\mathbf{g} = \mathbf{f} \circ \varphi: S' \to \mathbb{R}^3$ de K.

1. $\varphi: S' = \{(s, t) | s^2 + t^2 \leq 1\} \to \mathbb{R}^2$, $\varphi(s, t) = (-t, s)$.

2. $\varphi: S' = \{(s, t) | s^2 + t^2 \leq 1\} \to \mathbb{R}^2$, $\varphi(s, t) = (-s, -t)$.

3. $\varphi: S' = \{(s, t) | s^2 + t^2 \leq 4\} \to \mathbb{R}^2$, $\varphi(s, t) = (t/2, s/2)$.

4. $\varphi: S' = \{(s, t) | s^2 + t^2 \leq 1/2\} \to \mathbb{R}^2$, $\varphi(s, t) = (s, t)$.

5. $\varphi: S' = \{(s, t) | 2s^2 + 3t^2 \leq 6\} \to \mathbb{R}^2$, $\varphi(s, t) = (-s/\sqrt{3}, -t/\sqrt{2})$.

6. $\varphi: S \to \mathbb{R}^2$, $\varphi(s, t) = \sqrt{2}/2(t + s, t - s)$.

7. $\varphi: S \to \mathbb{R}^2$, $\varphi(s, t) = (\sqrt{3}s/2 - t/2, s/2 + \sqrt{3}t/2)$.

Sea $K = \mathbf{f}(S)$ una superficie simple, parametrizada por la función $\mathbf{f}: S \to \mathbb{R}^3$, definida en la región $S = \{(u, v) | a \leq u \leq b, c \leq v \leq d\}$. Determine con cuáles de las funciones $\varphi: S' \subset \mathbb{R}^2 \to \mathbb{R}^2$ dadas en los ejercicios 8–12, se pueden obtener reparametrizaciones $\mathbf{g} = \mathbf{f} \circ \varphi: S' \to \mathbb{R}^3$ de K.

8. $\varphi: S \to \mathbb{R}^2$, $\varphi(s, t) = (-t, s)$.

9. $\varphi: S \to \mathbb{R}^2$, $\varphi(s, t) = (t, -s)$.

10. $\varphi: S' = \{(s, t) | a/2 \leq s \leq b/2, c/3 \leq t \leq d/3\} \to \mathbb{R}^2$, $\varphi(s, t) = (2s, 3t)$.

11. $\varphi: S' = \{(s, t) | a/2 \leq s \leq b/2, c/3 \leq t \leq d/3\} \to \mathbb{R}^2$, $\varphi(s, t) = (t, s)$.

12. $\varphi: S' = \{(s, t) | 0 \leq s \leq 1, 0 \leq t \leq 1\} \to \mathbb{R}^2$, $\varphi(s, t) = ((b - a)s + a, (d - c)t + c)$

8.3 Espacios tangentes, planos tangentes y vectores normales

Sea $K = \mathbf{f}(S)$ una superficie simple, imagen de la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Sea $\mathbf{p} \in \operatorname{Int} S$ y $\mathbf{q} = \mathbf{f}(\mathbf{p})$ (entonces $\mathbf{q} \in \operatorname{Int} K$). Sea $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ un camino de clase $\mathscr{C}^1$ cuya imagen está contenida en K (es decir, $\boldsymbol{\lambda}([a, b]) \subset K$, o bien, los puntos de *la curva* descrita por el camino $\boldsymbol{\lambda}$ son puntos de K) tal que para algún $c \in [a, b]$, $\boldsymbol{\lambda}(c) = \mathbf{q}$. O sea, la traza de $\boldsymbol{\lambda}$ es una curva en K que pasa por $\mathbf{q}$. El vector $\boldsymbol{\lambda}'(c)$ es un vector tangente a la curva en $\mathbf{q} = \boldsymbol{\lambda}(c)$, y, puesto que esta curva cobra vida en K, es natural que podamos decir también que $\boldsymbol{\lambda}'(c)$ es un vector *tangente a la superficie K en* $\mathbf{q}$. Se tiene, de hecho, la siguiente definición.

Definición. Sea $K = \mathbf{f}(S)$ una superficie simple, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Sea $\mathbf{q} \in \operatorname{Int} K$. Se dice que el vector $\mathbf{v} \in \mathbb{R}^3$ es un *vector tangente* a K en $\mathbf{q}$ si hay un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ de clase $\mathscr{C}^1$, de modo que $\boldsymbol{\lambda}([a, b]) \subset K$, $\boldsymbol{\lambda}(c) = \mathbf{q}$ para algún $c \in [a, b]$ y $\boldsymbol{\lambda}'(c) = \mathbf{v}$. Al conjunto de todos los vectores tangentes a K en $\mathbf{q}$, se le llama *espacio tangente a K en* $\mathbf{q}$, y se denota por $T_{\mathbf{q}}(K)$. ■

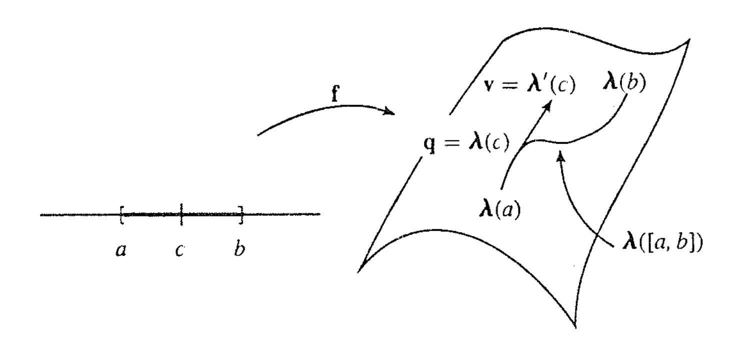


Figura 1. El camino $\boldsymbol{\lambda}$ y la superficie K.

El hecho más importante sobre el espacio tangente $T_{\mathbf{q}}(K)$ es que éste está íntimamente relacionado con los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$, derivadas parciales de la función $\mathbf{f}$ que parametriza a K, los cuales sabemos son linealmente independientes. Antes de enunciar el teorema correspondiente, recordemos que el *subespacio generado* por dos vectores $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$ es aquel conjunto de vectores $\mathbf{v} \in \mathbb{R}^3$ que se escriben como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2$ (ver capítulo 1 sección 1). Denotaremos este subespacio por $\mathscr{L}(\mathbf{v}_1, \mathbf{v}_2)$. Se tiene entonces que

$$\mathscr{L}(\mathbf{v}_1, \mathbf{v}_2) = \{\alpha \mathbf{v}_1 + \beta \mathbf{v}_2, \alpha, \beta \in \mathbb{R}\}$$

Teorema 8.3.1 Sea $K = \mathbf{f}(S)$ una superficie simple, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f} = (f_1, f_2, f_3)$. Sea $\mathbf{p} \in \operatorname{Int} S$, $\mathbf{q} = \mathbf{f}(\mathbf{p})$. Entonces

$$T_{\mathbf{q}}(K) = \mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right)$$ ■

Aunque no daremos todos los detalles que muestren la validez del resultado anterior, sí presentamos a continuación las ideas generales de su demostración.

Primeramente veamos que $\mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right) \subseteq T_{\mathbf{q}}(K)$. Para esto, tomemos α, β números reales arbitrarios y consideremos el vector $\mathbf{h} = (\alpha, \beta) \in \mathbb{R}^2$. Sea $\boldsymbol{\lambda}: [-\epsilon, \epsilon] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = \mathbf{f}(\mathbf{p}+t\mathbf{h})$. Como $\mathbf{p} \in \operatorname{Int} S$, con el número $\epsilon > 0$ suficientemente pequeño aseguramos que $\mathbf{p} + t\mathbf{h} \in S$. Obsérvese que éste es un camino de clase $\mathscr{C}^1$ (pues $\mathbf{f}$ es de clase $\mathscr{C}^1$) cuya imagen vive en K, de modo que $\boldsymbol{\lambda}(0) = \mathbf{f}(\mathbf{p}) = \mathbf{q}$.

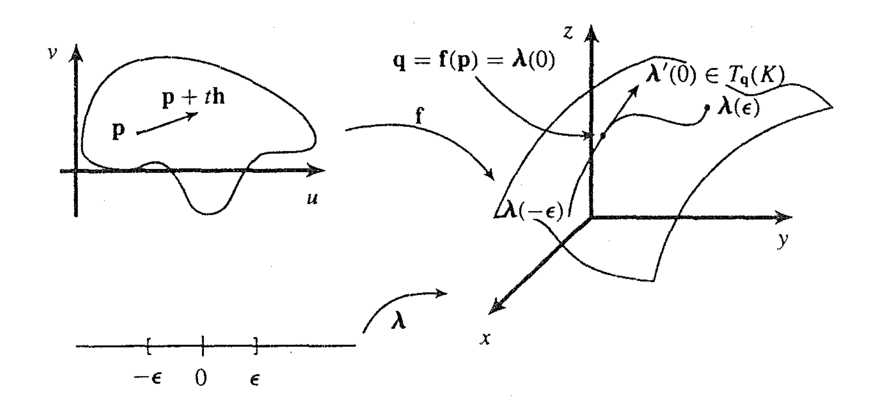


Figura 2. El camino $\boldsymbol{\lambda}(t) = \mathbf{f}(\mathbf{p} + t\mathbf{h})$.

Además, según la regla de la cadena, se tiene que

$$\underset{\text{vector en } \mathbb{R}^3}{\boldsymbol{\lambda}'(t)} = \underset{\text{matriz de } 3\times 2}{\mathbf{f}'(\mathbf{p} + t\mathbf{h})} \cdot \underset{\text{vector en } \mathbb{R}^2}{\mathbf{h}}$$

de modo que $\boldsymbol{\lambda}'(0) = \mathbf{f}'(\mathbf{p}) \cdot \mathbf{h}$, lo cual en un esquema se ve como

$$\boldsymbol{\lambda}'(0) = \begin{bmatrix} \uparrow & \uparrow \\ \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) & \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) + \beta \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$$

Como $\boldsymbol{\lambda}'(0) \in T_{\mathbf{q}}(K)$, vemos que cualquier combinación lineal de los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p})$, $\frac{\partial f}{\partial v}(\mathbf{p})$ es un vector de $T_{\mathbf{q}}(K)$. Esto muestra entonces la contención $\mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right) \subseteq T_{\mathbf{q}}(K)$.

Para ver la validez de la contención $T_{\mathbf{q}}(K) \subseteq \mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right)$, tomemos un vector $\mathbf{v} \in T_{\mathbf{q}}(K)$. Existe entonces un camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ de clase $\mathscr{C}^1$, $\boldsymbol{\lambda}([a, b]) \subset K$, $\boldsymbol{\lambda}(c) = \mathbf{q}$, para algún $c \in [a, b]$ y $\boldsymbol{\lambda}'(c) = \mathbf{v}$. Es posible demostrar (no lo haremos) que este camino $\boldsymbol{\lambda}$ es la composición de $\mathbf{f}$ con un camino $\boldsymbol{\mu}: [a, b] \to \mathbb{R}^2$ de clase $\mathscr{C}^1$, $\boldsymbol{\mu}([a, b]) \subset S$, $\boldsymbol{\mu}(c) = \mathbf{p}$, como se muestra en la figura 3.

Así pues, el camino $\boldsymbol{\lambda}$ es $\boldsymbol{\lambda} = \mathbf{f} \circ \boldsymbol{\mu}: [a, b] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (\mathbf{f} \circ \boldsymbol{\mu})(t) = \mathbf{f}(\boldsymbol{\mu}(t))$ y con $\boldsymbol{\lambda}(c) = \mathbf{f}(\boldsymbol{\mu}(c)) = \mathbf{f}(\mathbf{p}) = \mathbf{q}$. Según la regla de la cadena se tiene

$$\underset{\text{vector en } \mathbb{R}^3}{\boldsymbol{\lambda}'(t)} = \underset{\text{matriz de } 3\times 2}{\mathbf{f}'(\boldsymbol{\mu}(t))} \; \underset{\text{vector en } \mathbb{R}^2}{\boldsymbol{\mu}'(t)}$$

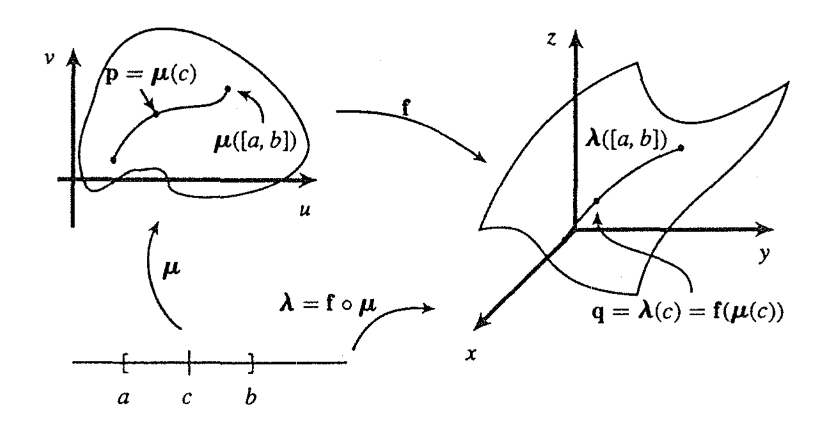


Figura 3. El camino $\boldsymbol{\lambda}$ como composición de $\mathbf{f}$ con $\boldsymbol{\mu}$

Si $\boldsymbol{\mu}'(c) = (a, b)$, se tiene

$$\mathbf{v} = \boldsymbol{\lambda}'(c) = \begin{bmatrix} \uparrow & \uparrow \\ \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) & \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = a\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) + b\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$$

de modo que el vector $\mathbf{v} \in T_{\mathbf{q}}(K)$ es una combinación lineal de los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$. Es decir $\mathbf{v} \in \mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right)$. Esto muestra la contención $T_{\mathbf{q}}(K) \subseteq \mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right)$.

Se tiene pues que el espacio tangente $T_{\mathbf{q}}(K)$ está constituido por todas las combinaciones lineales de los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p})$ y $\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$ donde $\mathbf{f}$ es una parametrización de K. Además, si $\mathbf{g} = \mathbf{f} \circ \boldsymbol{\varphi}: S' \subset \mathbb{R}^2 \to \mathbb{R}^3$ es una reparametrización de K, con $\mathbf{p}' \in S$, $\mathbf{g}(\mathbf{p}') = (\mathbf{f} \circ \boldsymbol{\varphi})(\mathbf{p}') = \mathbf{f}(\boldsymbol{\varphi}(\mathbf{p}')) = \mathbf{f}(\boldsymbol{\varphi}(\mathbf{p})) = \mathbf{q}$, usando las expresiones (A) de la sección anterior, con $(s, t) = \mathbf{p}'$, $\boldsymbol{\varphi}(\mathbf{p}') = \mathbf{p}$, se tiene

$$\frac{\partial \mathbf{g}}{\partial s}(\mathbf{p}') = a(\mathbf{p}')\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) + c(\mathbf{p}')\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$$
$$\frac{\partial \mathbf{g}}{\partial t}(\mathbf{p}') = b(\mathbf{p}')\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) + d(\mathbf{p}')\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$$

de donde, puesto que $a(\mathbf{p}')d(\mathbf{p}') - c(\mathbf{p}')b(\mathbf{p}') \neq 0$ (¿por qué?), se ve que

$$\mathscr{L}\left(\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}), \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})\right) = \mathscr{L}\left(\frac{\partial \mathbf{g}}{\partial s}(\mathbf{p}'), \frac{\partial \mathbf{g}}{\partial t}(\mathbf{p}')\right)$$

(¡verifique!). Así pues, el espacio tangente $T_{\mathbf{q}}(K)$ es el espacio generado por los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p})$ y $\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$ donde $\mathbf{f}$ es *cualquier* parametrización de K. Más aún, puesto que los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p})$ y $\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$ son linealmente independientes, podemos decir que estos constituyen una *base* del espacio $T_{\mathbf{q}}(K)$. Este es pues un subespacio de dimensión 2 en $\mathbb{R}^3$, el cual geométricamente es un *plano en* $\mathbb{R}^3$ *que pasa por el origen*.

Ejemplo 1. Consideremos la superficie $K = \mathbf{f}(S)$ del ejemplo 1 de la sección 1, parametrizada por $\mathbf{f}: [a, b] \times [c, d] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$. Sea $\mathbf{p} \in [a, b] \times [c, d]$ y $\mathbf{q} = \mathbf{f}(\mathbf{p})$. El

espacio tangente $T_{\mathbf{q}}(K)$ es el espacio generado por los vectores

$$\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) = (1, 0, \alpha) \quad \text{y} \quad \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = (0, 1, \beta)$$

Nótese que estos son vectores constantes (no dependen del punto $\mathbf{p}$), de modo que $T_{\mathbf{q}}(K)$ es el mismo para todo $\mathbf{q} \in K$. Se tiene entonces que

$$\begin{aligned} T_{\mathbf{q}}(K) &= \mathscr{L}((1, 0, \alpha), (0, 1, \beta)) \\ &= \{(x, y, z)|(x, y, z) = c_1(1, 0, \alpha) + c_2(0, 1, \beta), c_1, c_2 \in \mathbb{R}\} \\ &= \{(x, y, z)|z = \alpha x + \beta y\} \end{aligned}$$

Este espacio es en efecto un plano que pasa por el origen. Es, de hecho, un plano paralelo a K que pasa por el origen.

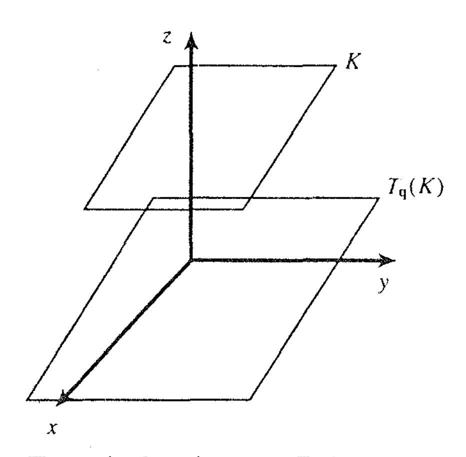


Figura 4. Los planos K y $T_{\mathbf{q}}(K)$. ■

Ejemplo 2. Consideremos la superficie parabólica $K = \mathbf{f}(S)$ del ejemplo 2 de la sección 1, imagen de $\mathbf{f}: S = \{(u, v)|u^2 + v^2 \leq 1\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$. Tomemos el punto $\mathbf{p} = (1/2, 1/2) \in \text{Int}\, S$. Entonces $\mathbf{q} = \mathbf{f}(\mathbf{p}) = \mathbf{f}(1/2, 1/2) = (1/2, 1/2, 1/2)$. El espacio tangente $T_{\mathbf{q}}(K)$ es el espacio generado por los vectores

$$\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) = (1, 0, 2u)\Big|_{\substack{u=1/2\\v=1/2}} = (1, 0, 1)$$

$$\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = (0, 1, 2v)\Big|_{\substack{u=1/2\\v=1/2}} = (0, 1, 1)$$

Es decir

$$\begin{aligned} T_{\mathbf{q}}(K) &= \{(x, y, z)|(x, y, z) = c_1(1, 0, 1) + c_2(0, 1, 1), c_1, c_2 \in \mathbb{R}\} \\ &= \{(x, y, z)|z = x + y\} \end{aligned}$$

Así pues, $T_{\mathbf{q}}(K)$ es el plano $z = x + y$. ■

En el capítulo 2 se presentó la idea de "plano tangente" a una superficie en un punto de ella. Es claro que la idea desarrollada en esta sección del *espacio tangente*, no es la del plano tangente mencionado, pues si bien el espacio tangente a una superficie es un plano, éste pasa por el origen y no podemos esperar, en general, que haga tangencia con la superficie en el punto considerado. Sin embargo, estando constituído el espacio $T_{\mathbf{q}}(K)$ por *vectores tangentes* a la superficie K en $\mathbf{q}$, todo lo que tenemos que hacer, para lograr establecer la idea de "plano tangente a K en $\mathbf{q}$", es *trasladar al punto* $\mathbf{q}$ *el espacio* $T_{\mathbf{q}}(K)$ *de vectores tangentes* a K *en* $\mathbf{q}$. Así pues, si consideramos el conjunto de vectores $(x, y, z) \in \mathbb{R}^3$ que se escriben como la suma de $\mathbf{q} \in K$ más un vector en $T_{\mathbf{q}}(K)$, estos vectores constituirán *el plano tangente* a K *en* $\mathbf{q}$. Este es pues el plano

$$\mathbf{q} + T_{\mathbf{q}}(K) = \{(x, y, z) | (x, y, z) = \mathbf{q} + \mathbf{v}, \ \mathbf{v} \in T_{\mathbf{q}}(K)\}$$

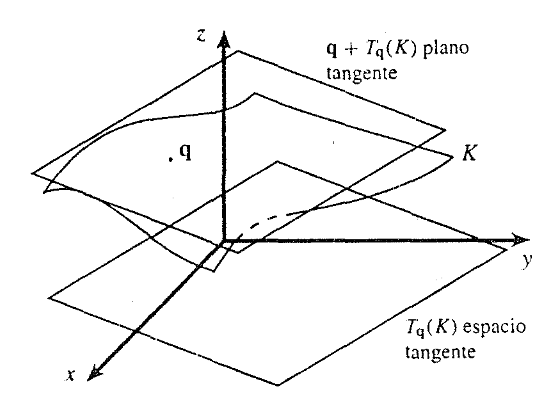


Figura 5. El plano tangente a K en $\mathbf{q}$.

Ejemplo 3. El plano tangente a la superficie parabólica K del ejemplo 2 en el punto $\mathbf{q} = (1/2, 1/2, 1/2)$ es

$$\begin{aligned}\mathbf{q} + T_{\mathbf{q}}(K) &= (1/2, 1/2, 1/2) + (x, y, x + y) = (x + 1/2, y + 1/2, x + y + 1/2) \\ &= (x, y, x + y - 1/2)\end{aligned}$$

Es decir, el plano tangente a la gráfica de la función $z = x^2 + y^2$ en el punto $\mathbf{q} = (1/2, 1/2, 1/2)$ es $z = x + y - 1/2$. ∎

Puesto que los vectores $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p})$ y $\frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$ constituyen una base del espacio tangente $T_{\mathbf{q}}(K)$ ($\mathbf{q} = \mathbf{f}(\mathbf{p})$), siendo este espacio un plano en $\mathbb{R}^3$, el producto cruz $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$ será un vector normal a este plano. Más aún, será un vector normal al plano tangente a la superficie K en $\mathbf{q}$. Así pues podemos decir que el plano tangente a K en $\mathbf{q}$ tiene por ecuación

$$\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) \cdot ((x, y, z) - \mathbf{q}) = 0$$

Ejemplo 4. De nuevo, la superficie parabólica del ejemplo 2 considerada en el ejemplo anterior. En el punto $\mathbf{p} = (1/2, 1/2)$ se tiene

$$\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) = (1, 0, 1) \qquad \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = (0, 1, 1)$$

de modo que un vector normal al plano tangente a la superficie en $\mathbf{q} = \mathbf{f}(\mathbf{p}) = (1/2, 1/2, 1/2)$ es

$$\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = (1, 0, 1) \times (0, 1, 1) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = (-1, -1, 1)$$

y entonces el plano tangente mencionado tiene por ecuación a

$$(-1, -1, 1) \cdot ((x, y, z) - (1/2, 1/2, 1/2)) = 0$$

o sea $-(x - 1/2) - (y - 1/2) + (z - 1/2) = 0$, o bien $z = x + y - 1/2$ como se obtuvo en el ejemplo anterior. ■

Denotemos por $N_{\mathbf{f}}(\mathbf{p})$ al vector $\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$, el cual es, como sabemos, normal al plano tangente a la superficie K en $\mathbf{q} = \mathbf{f}(\mathbf{p})$. Si $\mathbf{g} = \mathbf{f} \circ \varphi$ es una reparametrización de K, habíamos ya observado que los vectores $\frac{\partial \mathbf{g}}{\partial s}(\mathbf{p}')$ y $\frac{\partial \mathbf{g}}{\partial t}(\mathbf{p}')$, $\varphi(\mathbf{p}') = \mathbf{p}$, constituían también una base del espacio tangente $T_{\mathbf{q}}(K)$, de modo que el vector $\frac{\partial \mathbf{g}}{\partial s}(\mathbf{p}') \times \frac{\partial \mathbf{g}}{\partial t}(\mathbf{p}') = N_{\mathbf{g}}(\mathbf{p}')$ también debería ser un vector normal al plano tangente a K en $\mathbf{q}$. Los vectores $N_{\mathbf{f}}(\mathbf{p})$ y $N_{\mathbf{g}}(\mathbf{p}')$ deben ser entonces linealmente dependientes. Es decir, uno de ellos debe ser un múltiplo del otro. La relación precisa entre estos dos vectores ha quedado plenamente establecida en la sección 2, donde obtuvimos (fórmula (Â)) que

$$\frac{\partial \mathbf{g}}{\partial s}(s, t) \times \frac{\partial \mathbf{g}}{\partial t}(s, t) = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} \frac{\partial \mathbf{f}}{\partial u}(\varphi(s, t)) \times \frac{\partial \mathbf{f}}{\partial v}(\varphi(s, t))$$

o bien, poniendo $\mathbf{p}' = (s, t)$, $\mathbf{p} = (u, v) = \varphi(s, t)$

$$N_{\mathbf{g}}(\mathbf{p}') = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)}(\mathbf{p}') N_{\mathbf{f}}(\mathbf{p})$$

donde $\mathbf{q} = \mathbf{f}(\mathbf{p}) = \mathbf{f}(\varphi(\mathbf{p}'))$. Esquemáticamente

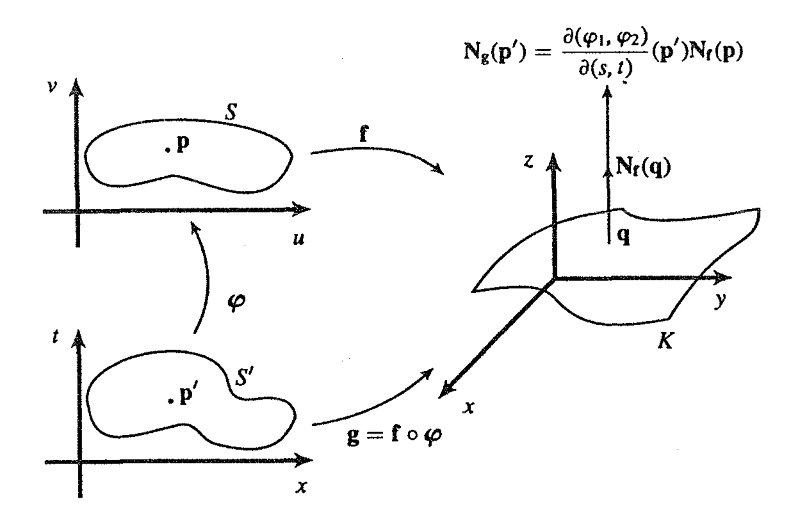


Figura 6. Los vectores $N_{\mathbf{f}}$ y $N_{\mathbf{g}}$.

Ejemplo 5. Consideremos la superficie parabólica $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: \{(u, v)|u^2 + v^2 \leq 10\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$. Sea $\mathbf{g}: \{(s, t)|s^2 + t^2 \leq 5\} \to \mathbb{R}^3$ la función $\mathbf{g}(s, t) = (s+t, s-t, 2s^2+2t^2)$. Esta es una reparametrización de K. De hecho, si $\varphi: \{(s, t)|s^2+t^2 \leq 5\} \to \{(u, v)|u^2 + v^2 \leq 10\}$ es $\varphi(s, t) = (s + t, s - t)$, la cual es una biyección de clase $\mathscr{C}^1$ tal que

$$\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} = \det \begin{bmatrix} \dfrac{\partial\varphi_1}{\partial s} & \dfrac{\partial\varphi_1}{\partial t} \\ \dfrac{\partial\varphi_2}{\partial s} & \dfrac{\partial\varphi_2}{\partial t} \end{bmatrix} = \det \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = -2 \neq 0$$

se tiene que $\mathbf{g}(s, t) = (\mathbf{f} \circ \varphi)(s, t) = \mathbf{f}(s + t, s - t) = (s + t, s - t, 2s^2 + 2t^2)$. Tomemos el punto $\mathbf{p} = (1, 1) \in S = \{(u, v)|u^2 + v^2 \leq 10\}$. El punto $\mathbf{p}' = (1, 0) \in S' = \{(s, t)|s^2 + t^2 \leq 5\}$ es tal que $\mathbf{q} = (1, 1, 2) = \mathbf{f}(\mathbf{p}) = \mathbf{g}(\mathbf{p}')$. Se tiene

$$\frac{\partial \mathbf{f}}{\partial u} = (1, 0, 2u) \qquad \frac{\partial \mathbf{f}}{\partial v} = (0, 1, 2v)$$
$$\frac{\partial \mathbf{g}}{\partial s} = (1, 1, 4s) \qquad \frac{\partial \mathbf{g}}{\partial t} = (1, -1, 4t)$$

de donde

$$\frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) = (1, 0, 2), \ \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = (0, 1, 2), \ \frac{\partial \mathbf{g}}{\partial s}(\mathbf{p}') = (1, 1, 4), \ \frac{\partial \mathbf{g}}{\partial t}(\mathbf{p}') = (1, -1, 0)$$

Entonces

$$\begin{aligned} N_{\mathbf{f}}(\mathbf{p}) &= \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = (1, 0, 2) \times (0, 1, 2) \\ &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = (-2, -2, 1) \\ N_{\mathbf{g}}(\mathbf{p}') &= \frac{\partial \mathbf{g}}{\partial s}(\mathbf{p}') \times \frac{\partial \mathbf{g}}{\partial t}(\mathbf{p}') = (1, 1, 4) \times (1, -1, 0) \\ &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 4 \\ 1 & -1 & 0 \end{bmatrix} = (4, 4, -2) \end{aligned}$$

Se tiene entonces que

$$\begin{array}{ccc} (4, 4, -2) = & -2 & (-2, -2, 1) \\ \downarrow & \downarrow & \downarrow \\ N_{\mathbf{g}}(\mathbf{p}') & \frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)}(\mathbf{p}') & N_{\mathbf{f}}(\mathbf{p}') \end{array}$$

como debía ocurrir. ■

Ejemplo 6. Sea $\varphi: S \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathscr{C}^1$ definida en la región simple del tipo I y II, S. Sabemos que la gráfica de φ es una superficie simple parametrizada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \varphi(u, v))$.

En este caso el vector normal $N_{\mathbf{f}}(u, v)$ se ve como

$$N_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}) = \left(1, 0, \frac{\partial \varphi}{\partial u}\right) \times \left(0, 1, \frac{\partial \varphi}{\partial v}\right)$$

$$= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial \varphi}{\partial u} \\ 0 & 1 & \frac{\partial \varphi}{\partial v} \end{bmatrix} = \left(-\frac{\partial \varphi}{\partial u}, -\frac{\partial \varphi}{\partial v}, 1\right)$$

Este resultado lo conocíamos ya desde el capítulo 2. ∎

Ejercicios (Capítulo 8, Sección 3)

1. Considere la superficie simple K, gráfica de la función de clase $\mathscr{C}^1$, $z = \varphi(x, y)$, digamos que parametrizada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \varphi(u, v))$. Demuestre que el espacio tangente a K en el punto $\mathbf{q} = (x_0, y_0, \varphi(x_0, y_0)) \in \operatorname{Int} K$ es

$$T_{\mathbf{q}}(K) = \{(x, y, z) | z = \frac{\partial \varphi}{\partial x}(x_0, y_0)x + \frac{\partial \varphi}{\partial y}(x_0, y_0)y\}$$

y que el plano tangente a K en $\mathbf{q}$ es

$$\mathbf{q} + T_{\mathbf{q}}(K) = \{(x, y, z) | z = \frac{\partial \varphi}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial \varphi}{\partial y}(x_0, y_0)(y - y_0) + \varphi(x_0, y_0)\}$$

(Este último resultado lo conocíamos ya desde el capítulo 2).

2. Considere la superficie simple K parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Suponga que en los alrededores del punto $\mathbf{q} = (x_0, y_0, z_0) = \mathbf{f}(\mathbf{p}) \in \operatorname{Int} K$, la superficie K se ve como un nivel constante de la función de clase $\mathscr{C}^1$, $h: U \subseteq \mathbb{R}^3 \to \mathbb{R}$. Demuestre que el espacio tangente a K en $\mathbf{q}$ es

$$T_{\mathbf{q}}(K) = \{(x, y, z) | \frac{\partial h}{\partial x}(\mathbf{q})x + \frac{\partial h}{\partial y}(\mathbf{q})y + \frac{\partial h}{\partial z}(\mathbf{q})z = 0\}$$

y que el plano tangente a K en $\mathbf{q}$ es

$$\mathbf{q} + T_{\mathbf{q}}(K) = \{(x, y, z) | \frac{\partial h}{\partial x}(\mathbf{q})(x - x_0) + \frac{\partial h}{\partial y}(\mathbf{q})(y - y_0) + \frac{\partial h}{\partial z}(\mathbf{q})(z - z_0) = 0\}$$

(Este resultado lo conocíamos también del capítulo 2).

En los ejercicios 3–10, determine el espacio tangente y el plano tangente a la superficie simple K en el punto $\mathbf{q}$ indicado.

3. $K = \mathbf{f}(S)$, donde $\mathbf{f}: S = [-2, 2] \times [-2, 2] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + 4v^2)$, en el punto $\mathbf{q} = \mathbf{f}(1, 1)$.
4. $K = \mathbf{f}(S)$, donde $\mathbf{f}: S = [0, 3] \times [0, 3] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, 3u^2 + v^3, 2v)$, en el punto $\mathbf{q} = \mathbf{f}(1, 2)$.

5. $K = \mathbf{f}(S)$, donde $\mathbf{f}: S = \{(u, v)|u^2 + v^2 \leq 4\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (3uv + v^2, v, 4u)$, en el punto $\mathbf{q} = \mathbf{f}(-1, -1)$.

6. $K = \{(x, y, z)|x^2 + y^2 + z^2 = 14, z \geq 0\}$, en el punto $\mathbf{q} = (1, 2, 3)$.

7. $K = \{(x, y, z)|2x^2 + y^2 + z^2 = 2, 1/3 \leq x \leq 1\}$, en el punto $\mathbf{q} = (1, 0, 0)$.

8. $K = \{(x, y, z)|z = 3x^2 + 5y^2, x^2 + y^2 \leq 4x\}$, en el punto $\mathbf{q} = (1, 0, 3)$.

9. $K = \{(x, y, z)|y = x^2 + z^2 - 2x + 4z - 6, 0 \leq x \leq 3, 0 \leq z \leq 3 - x\}$, en el punto $\mathbf{q} = (1, -2, 1)$.

10. $K = \{(x, y, z)|z = e^{-(2x^2+4y^2)}, x^2 + y^2 \leq 2x\}$, en el punto $\mathbf{q} = (1, 0, e^{-2})$.

11. Considere la función de clase $\mathscr{C}^1$, $\varphi: \mathbb{R}^2 \to \mathbb{R}$. Suponga que esta función tiene un extremo local en el punto $\mathbf{p} = (x_0, y_0)$. Considere la superficie simple $K = \mathbf{f}(S)$, parametrizada por la función $\mathbf{f}: S \to \mathbb{R}^3$, definida en la región $S = \{(u, v)|(u - x_0)^2 + (v - y_0)^2 \leq 1\}$, dada por:
   a. $\mathbf{f}(u, v) = (u, v, \varphi(u, v))$,
   b. $\mathbf{f}(u, v) = (u, \varphi(u, v), v)$,
   c. $\mathbf{f}(u, v) = (\varphi(u, v), u, v)$.

Determine en cada caso el espacio tangente y el plano tangente a K en el punto $\mathbf{q} = \mathbf{f}(\mathbf{p})$.

12. Sea $K = \mathbf{f}(S)$ la superficie simple parametrizada por $\mathbf{f}: S = \{(u, v)|u^2 + v^2 \leq 1\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, 3uv + 2u + 4v, v)$. Sea $\varphi: S \to S$, $\varphi(s, t) = \sqrt{2}/2(s + t, s - t)$.
   a. Verifique que $\mathbf{g} = \mathbf{f} \circ \varphi: S \to \mathbb{R}^3$ es una reparametrización de K.
   b. Verifique que

$$\frac{\partial \mathbf{g}}{\partial s}(s, t) \times \frac{\partial \mathbf{g}}{\partial t}(s, t) = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} \frac{\partial \mathbf{f}}{\partial u}(\varphi(s, t)) \times \frac{\partial \mathbf{f}}{\partial v}(\varphi(s, t))$$

   c. Determine el espacio tangente y el plano tangente a K en el punto $\mathbf{q} = (1/2, 7/12, 1/3)$

8.4 Superficies mas generales

Con el concepto de superficie simple, dado en la sección 1, y con el que hemos venido trabajando hasta este momento, es posible definir ideas importantes y establecer hechos relevantes en la teoría de integrales de superficie que emprenderemos en el próximo capítulo, para el cual importa de modo particular el estudio del presente capítulo. Sin embargo, las superficies simples nos limitan a no poder considerar (desde esa óptica) un cubo o una esfera como una de tales superficies. En esta sección ampliaremos el concepto de superficie simple para definir algo que llamaremos simplemente "superficie", y que nos permitirá establecer algunos otros hechos importantes en el capítulo siguiente.

La idea general que habrá detrás del concepto de superficie que daremos, es que este tipo de objetos matemáticos se construyen "pegando" superficies simples. Consideremos, a manera de ejemplo introductorio, la siguiente situación:

Sea $K_1 = \mathbf{f}_1(S_1)$, la superficie simple imagen de la función $\mathbf{f}_1: S_1 = [0, 1] \times [0, 1] \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}_1(u, v) = (u, v, 1)$. K_1 es entonces el pedazo del plano $z = 1$ que se encuentra sobre la región $S_1 = [0, 1] \times [0, 1]$. Sea $K_2 = \mathbf{f}_2(S_2)$, la superficie simple imagen de la función $\mathbf{f}_2: S_2 = [0, 1] \times [1, 2] \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}_2(u, v) = (u, 1, v)$. K_2 es entonces el pedazo del plano $y = 1$

cuya proyección al plano xz es $S_2 = [0, 1] \times [1, 2]$. Definamos la función $\mathbf{f}: S = S_1 \cup S_2 \subset \mathbb{R}^2 \to \mathbb{R}^3$ como

$$\mathbf{f}(u, v) = \begin{cases} \mathbf{f}_1(u, v) & \text{si } (u, v) \in S_1 \\ \mathbf{f}_2(u, v) & \text{si } (u, v) \in S_2 \end{cases}$$

Centremos nuestra atención a los puntos interiores de S_1 y en los de S_2. Llame $U = \operatorname{Int} S_1 \cup \operatorname{Int} S_2$, es decir, $U = (0, 1) \times (0, 1) \cup (0, 1) \times (1, 2)$. Obsérvese que la función $\mathbf{f}$ es perfectamente bien portada en U: es de clase $\mathscr{C}^1$, es inyectiva y los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ son linealmente independientes. En forma geométrica se tiene

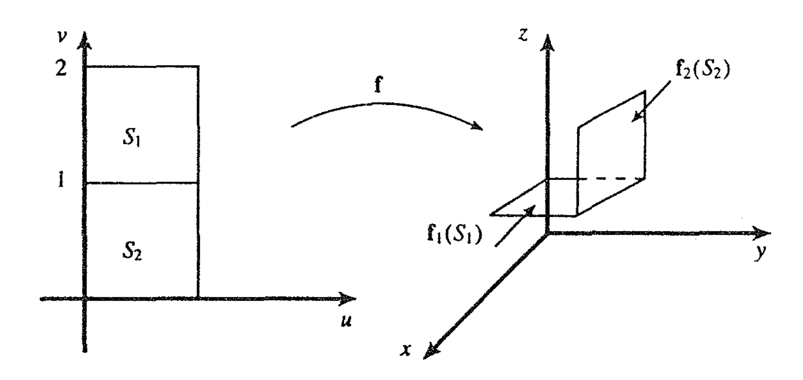


Figura 1. Superficie "unión" de las superficies simples $\mathbf{f}_1(S_1)$ y $\mathbf{f}_2(S_2)$.

A este tipo de objetos, como $K = \mathbf{f}(S)$, que se obtienen al pegar superficies simples las llamaremos "superficies". Por supuesto que el "pegado" se debe hacer con algunos cuidados. Estos aparecen en la siguiente definición.

Definición. Sean $S_1, S_2, \ldots, S_k$ regiones del tipo I y II en $\mathbb{R}^2$ tales que sus interiores son disjuntos por pares (es decir, $\operatorname{Int} S_i \cap \operatorname{Int} S_j = \varnothing$, $i \neq j$). Sea $S = S_1 \cup S_2 \cup \ldots \cup S_k$, $F = \partial S_1 \cup \partial S_2 \cup \ldots \cup \partial S_k$, $U = \operatorname{Int} S_1 \cup \operatorname{Int} S_2 \cup \ldots \cup \operatorname{Int} S_k$. Sea $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ una función de clase $\mathscr{C}^1$ tal que:

a. $\mathbf{f}: U \subset \mathbb{R}^2 \to \mathbb{R}^3$ es inyectiva.

b. los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ son linealmente independientes para todo (u, v) en U.

c. $\mathbf{f}(U) \cap \mathbf{f}(F) = \varnothing$

Entonces a $K = \mathbf{f}(S)$ se le llama *superficie* y a $\mathbf{f}$ se le dice ser una parametrización de K. ■

Así pues, una superficie es la imagen bajo una función $\mathbf{f}$ de una unión de regiones simples S_i del tipo I y II en $\mathbb{R}^2$ con interiores disjuntos. La función $\mathbf{f}$ debe ser de clase $\mathscr{C}^1$ y tal que restringida a la unión de los interiores de S_i sea inyectiva y tenga a los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ linealmente independientes. Además, se pide que la imagen de la unión de los interiores de S_i no compartan puntos con la imagen de la unión de sus fronteras. Esta última condición garantiza que las imágenes de cada una de las regiones S_i no se "traslapen".

Nótese entonces que las propiedades de $\mathbf{f}$ de ser inyectiva, de clase $\mathscr{C}^1$ y con los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ linealmente independientes, se satisfacen *casi en toda* su imagen, pues la imagen de la unión de las fronteras de las regiones (que es donde *no* se pide que se cumplan las propiedades citadas de $\mathbf{f}$) son "curvas" en K, las cuales, se puede demostrar, representan conjuntos de medida cero.

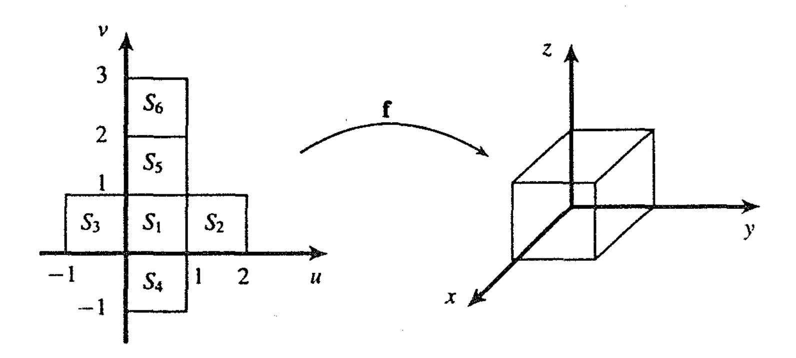


Figura 2. El cubo como superficie.

El resto de la presente sección lo dedicaremos a presentar algunos ejemplos de superficies.

Ejemplo 1. (Un cubo es una superficie). La idea geométrica que perseguimos en este ejemplo está inspirada en la figura 2.
Las 6 regiones mostradas cumplen con las condiciones de la definición: son del tipo I y II y sus interiores son disjuntos por pares. Cada una de las regiones S_i "será llevada por **f**" a una de las caras del cubo K. Obsérvese, de hecho, que cada cara del cubo es una superficie simple. Sus correspondientes parametrizaciones son

$$
\begin{array}{lll}
S_1 = [0,1] \times [0,1], & \mathbf{f}_1 \colon S_1 \to \mathbb{R}^3, & \mathbf{f}_1(u,v) = (u,v,0) \\
S_2 = [1,2] \times [0,1], & \mathbf{f}_2 \colon S_2 \to \mathbb{R}^3, & \mathbf{f}_2(u,v) = (1,u-1,v) \\
S_3 = [-1,0] \times [0,1], & \mathbf{f}_3 \colon S_3 \to \mathbb{R}^3, & \mathbf{f}_3(u,v) = (0,u+1,v) \\
S_4 = [0,1] \times [-1,0], & \mathbf{f}_4 \colon S_4 \to \mathbb{R}^3, & \mathbf{f}_4(u,v) = (u,0,v+1) \\
S_5 = [0,1] \times [1,2], & \mathbf{f}_5 \colon S_5 \to \mathbb{R}^3, & \mathbf{f}_5(u,v) = (u,1,v-1) \\
S_6 = [0,1] \times [2,3], & \mathbf{f}_6 \colon S_6 \to \mathbb{R}^3, & \mathbf{f}_6(u,v) = (u,v-2,1)
\end{array}
$$

Se tiene que

$$
\begin{aligned}
K_1 &= \mathbf{f}_1(S_1) \text{ es el piso de } K \\
K_2 &= \mathbf{f}_2(S_2) \text{ es el lado del frente de } K \\
K_3 &= \mathbf{f}_3(S_3) \text{ es el lado de atrás de } K \\
K_4 &= \mathbf{f}_4(S_4) \text{ es el lado izquierdo de } K \\
K_5 &= \mathbf{f}_5(S_5) \text{ es el lado derecho de } K \\
K_6 &= \mathbf{f}_6(S_6) \text{ es el techo de } K
\end{aligned}
$$

La función $\mathbf{f} \colon S \subset \mathbb{R}^2 \to \mathbb{R}^3$ definida en $S = \cup_{i=1}^{6} S_i$ como $\mathbf{f}(u,v) = \mathbf{f}_i(u,v)$ si $(u,v) \in S_i$, $i = 1, 2, \ldots, 6$, cumple con todas las propiedades establecidas en la definición dada anteriormente. El cubo K es entonces una superficie y $\mathbf{f}$ es una parametrización de él.

Ejemplo 2. (La esfera como una superficie). En el ejemplo 6 de la sección 1 se vio que la esfera $x^2 + y^2 + z^2 = r^2$ no puede ser considerada globalmente como una superficie simple. Sin embargo,

cada una de las semiesferas superior e inferior, sí lo es, y la proyección estereográfica proporcionó una parametrización de estos hemisferios. Podemos pensar entonces a la esfera completa como "la unión" de sus hemisferios norte y sur, coincidiendo estos en el ecuador (en esta línea la función **f** que parametriza a la esfera no es inyectiva)

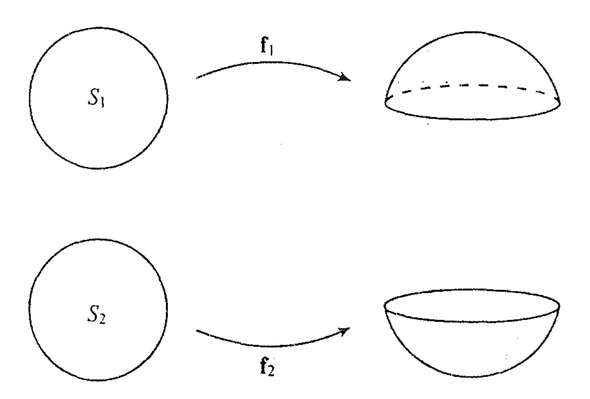


Figura 3. La esfera como superficie.

Otra manera de ver a la esfera completa como una superficie, es por medio de la parametrización dada por las coordenadas esféricas: $\mathbf{f}: [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = (r \cos u \operatorname{sen} v, r \operatorname{sen} u \operatorname{sen} v, r \cos v)$$

Esta función no es inyectiva, pues "encima" los puntos de la frontera ∂S a la misma imagen, el semicírculo $x^2 + z^2 = r^2$, $x \geq 0$, $y = 0$. En forma más precisa, observe que $\mathbf{f}(0, v) = \mathbf{f}(2\pi, v) = (r \operatorname{sen} v, 0, r \cos v)$ y $\mathbf{f}(u, 0) = (0, 0, r)$, $\mathbf{f}(u, \pi) = (0, 0, -r)$. Esta situación está considerada en la definición dada de superficie. Se verifica que todas las demás propiedades requeridas para **f** se satisfacen. ∎

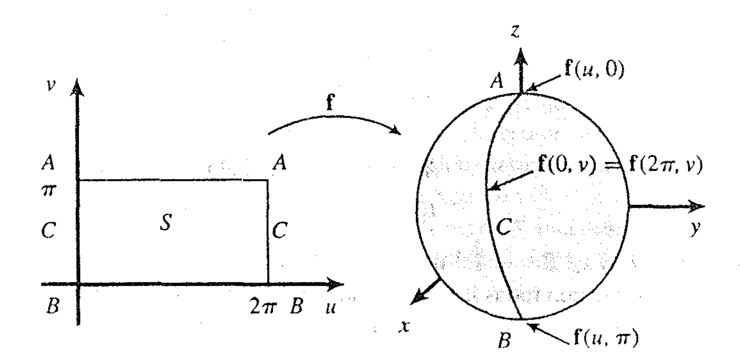


Figura 4. La función **f** no es inyectiva en la frontera de S.

Ejemplo 3. (El cono es una superficie). Sea $\mathbf{f}$ la función $\mathbf{f}: S = [0, 2\pi] \times [0, h] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (v \cos u, v \operatorname{sen} u, v)$. La imagen $K = \mathbf{f}(S)$ es un semicono con vértice en el origen y altura h, gráfica de la función $z = \sqrt{x^2 + y^2}$, $z \leq h$ (en esta función hemos puesto $x = v \cos u$, $y = v \operatorname{sen} u$, $z = v$). Obsérvese que $\mathbf{f}(0, v) = \mathbf{f}(2\pi, v) = (v, 0, v)$ y $\mathbf{f}(u, 0) = (0, 0, 0)$, de modo que $\mathbf{f}$ no es inyectiva en la frontera de S. Más aún, como

$$\frac{\partial \mathbf{f}}{\partial u} = (-v \operatorname{sen} u, v \cos u, 0), \qquad \frac{\partial \mathbf{f}}{\partial v} = (\cos u, \operatorname{sen} u, 1)$$

se tiene que en los puntos $(u, 0) \in S$ (que son mandados por $\mathbf{f}$ al origen de $\mathbb{R}^3$, vértice del cono), los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ son linealmente dependientes ($\frac{\partial \mathbf{f}}{\partial u}$ es el vector cero). Así pues, $\mathbf{f}$ no tiene un "buen comportamiento" en la frontera de S. En el interior de S es claro que $\mathbf{f}$ cumple con los requisitos establecidos en la definición de superficie, de modo que $K = \mathbf{f}(S)$ es una superficie parametrizada por $\mathbf{f}$.

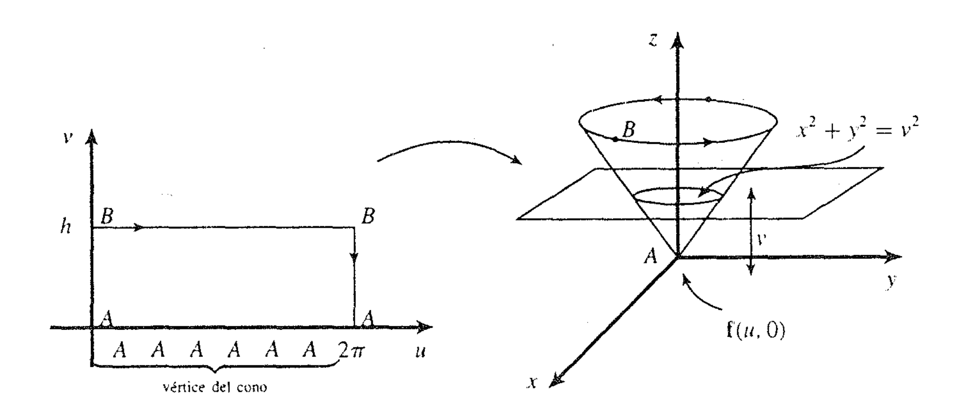


Figura 5. El cono no es una superficie simple, pero es una superficie. ■

Ejemplo 4. (El cilindro es una superficie). En el ejemplo 8 de la sección 1 se vio que el cilindro $x^2 + y^2 = 1$, $-1 \leq z \leq 1$ es una superficie seccionalmente simple. Podemos pensar también este cilindro como una superficie según la definición que estamos estudiando en esta sección, considerando la función $\mathbf{f}: S = [0, 2\pi] \times [-1, 1] \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = (\cos u, \operatorname{sen} u, v)$$

Dejamos al lector que se convenza de que $K = \mathbf{f}(S)$ es una superficie, verificando que $\mathbf{f}$ cumpla con todas las propiedades establecidas en la definición.

Ejemplo 5. (La banda de Möbius). En el ejemplo anterior la función $\mathbf{f}$ que parametriza el cilindro hace coincidir el lado $u = 0$, $-1 \leq v \leq 1$, con el lado $u = 2\pi$, $-1 \leq v \leq 1$, "enrollando" el rectángulo S. Una variante de esta situación es hacer que $\mathbf{f}$ haga coincidir los lados anteriormente mencionados, *pero invirtiendo uno de ellos*, de modo que el punto inferior del lado derecho coincida con el superior del lado izquierdo y el punto superior del lado derecho coincida con el inferior del lado izquierdo.

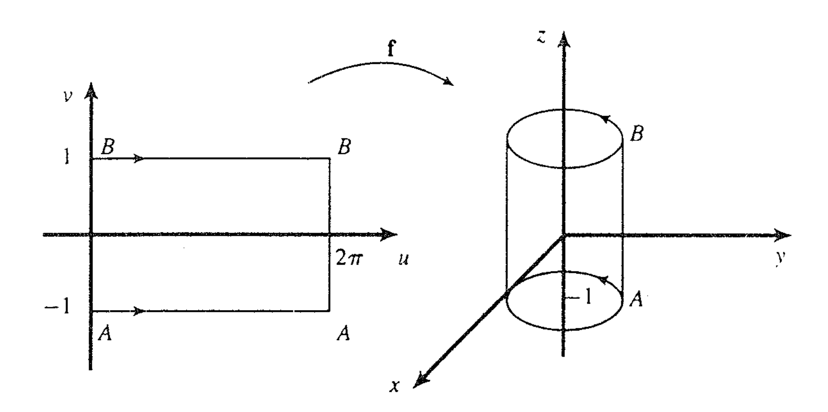


Figura 6. El cilindro como superficie.

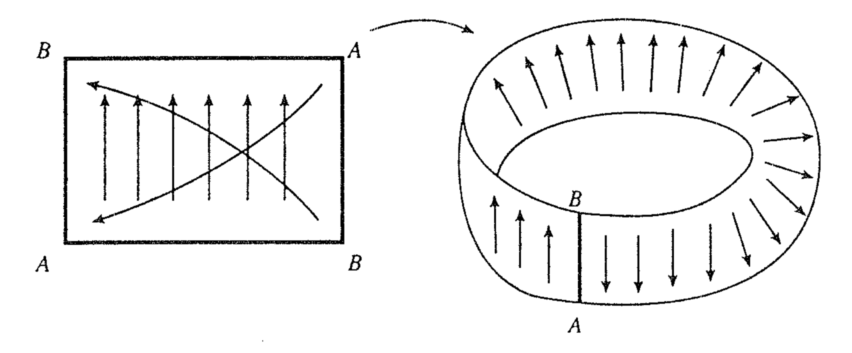


Figura 7. La banda de Möbius

Se obtiene así una interesante superficie conocida como "banda de Möbius". La función que parametriza a esta superficie K es $\mathbf{f}\colon [0, 2\pi] \times [-1, 1] \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = \left(\left(1 - v \operatorname{sen} \frac{u}{2}\right) \cos u, \left(1 - v \operatorname{sen} \frac{u}{2}\right) \operatorname{sen} u, v \cos \frac{u}{2}\right)$$

Nótese que

$$\mathbf{f}(A) = \mathbf{f}(0, -1) = (1, 0, -1) = \mathbf{f}(2\pi, 1) \quad \text{y} \quad \mathbf{f}(B) = \mathbf{f}(0, 1) = (1, 0, 1) = \mathbf{f}(2\pi, -1)$$

lo que confirma la "inversión" de los extremos de los lados del rectángulo S. ∎

Ejemplo 6. (El toro es una superficie). Tomemos un círculo C_1 de radio r digamos que en el semiplano $y > 0$, y hagámoslo girar alrededor del eje y. Llamemos R a la distancia del centro de C_1 al eje de rotación. Se tiene así un círculo "pequeño" C_1, cuyo centro gira en un plano perpendicular al plano que contiene a C_1, describiendo un círculo C_2 de radio R. La superficie generada por C_1 se

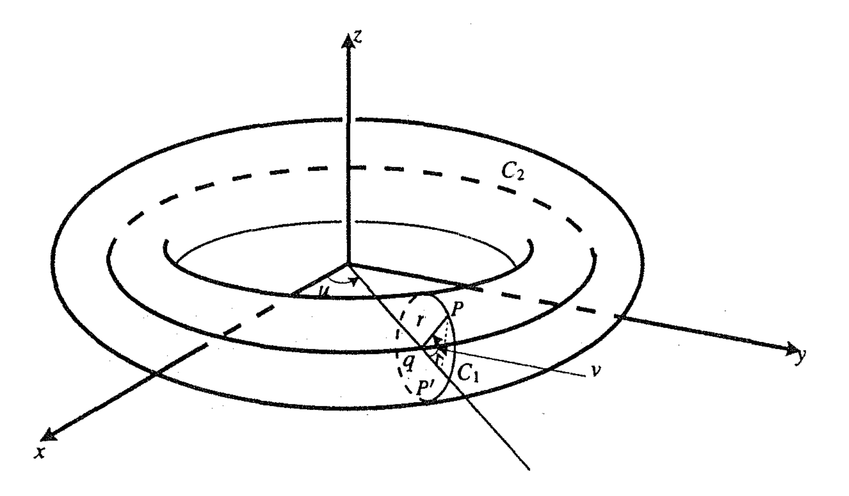


Figura 8. El toro.

le llama "toro" (la cual es simplemente una "dona"). Vamos a obtener una parametrización de ésta superficie. Designemos por u y v los ángulos mostrados en la figura 8.

Se trata de describir las coordenadas x, y, z del punto P en términos de los parámetros u y v. Se tiene la situación siguiente

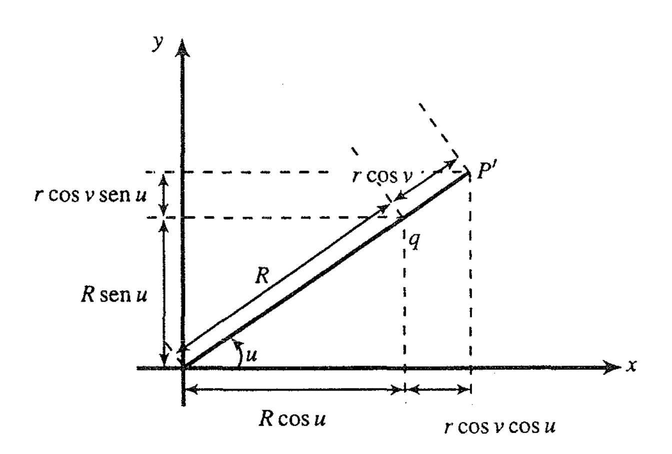


Figura 9. Las coordenadas x, y del punto P' en términos de u y v.

Las coordenadas del punto P', proyección de P en el plano xy, son entonces

$$x = R\cos u + r\cos v\cos u = (R + r\cos v)\cos u$$
$$y = R\operatorname{sen} u + r\cos v\operatorname{sen} u = (R + r\cos v)\operatorname{sen} u$$

Estas también son las coordenadas x, y del punto P. La coordenada z se obtiene fácilmente del triángulo $\mathbf{q}P'P$. Esta es

$$z = r \operatorname{sen} v$$

Así pues, la función $\mathbf{f}: [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$ dada por

$$\mathbf{f}(u, v) = ((R + r\cos v)\cos u, (R + r\cos v)\operatorname{sen} u, r \operatorname{sen} v)$$

es una parametrización del toro. ■

Es un hecho común que la función $\mathbf{f}$ que parametriza a una superficie K pierde la inyectividad en los puntos correspondientes a las curvas de K donde $\mathbf{f}$ "cierra" o "pega" partes distintas de la frontera de $S \subset \mathbb{R}^2$, la región donde $\mathbf{f}$ está definida. Esta situación ocurrió en *todas* las superficies de los 6 ejemplos anteriores (regrese a cada uno de ellos e identifique los "cierres" de S en K). Quisiéramos llamar la atención al hecho de que aún en los puntos de K que son imágenes de puntos distintos de la frontera de S puede ser posible determinar planos tangentes a la superficie: en realidad la posibilidad de poder hacerlo depende solamente de que el vector normal $N_{\mathbf{f}}(\mathbf{p}) = \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p})$ esté *bien determinado*. Esto ocurre, por supuesto, en los puntos del interior de S (¿por qué?), pero *puede también ocurrir en los puntos de la frontera de S*. Bastaría verificar que si $\mathbf{p}$ y $\mathbf{p}'$ son dos puntos de la frontera de S que tienen la misma imagen $\mathbf{q} = \mathbf{f}(\mathbf{p}) = \mathbf{f}(\mathbf{p}')$, los vectores

$$N_{\mathbf{f}}(\mathbf{p}) = \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}), \quad N_{\mathbf{f}}(\mathbf{p}') = \frac{\partial \mathbf{f}}{\partial u}(\mathbf{p}') \times \frac{\partial \mathbf{f}}{\partial v}(\mathbf{p}')$$

(que son no nulos) son linealmente dependientes (es decir, que sea uno un múltiplo del otro). De esta manera, tendríamos garantizada la existencia de un vector normal al plano tangente en $\mathbf{q}$ y así podríamos obtener su ecuación.

Ejemplo 7. Retomemos el cubo K del ejemplo 1.

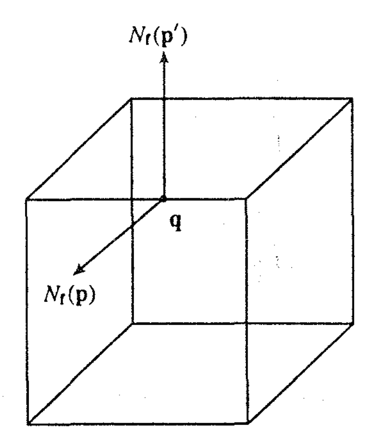


Figura 10. El cubo K

El punto $\mathbf{q} = (1, 1/2, 1)$ mostrado en la figura es imagen de los puntos $\mathbf{p} = (3/2, 1) \in S_2 = [1, 2] \times [0, 1]$ y $\mathbf{p}' = (1, 5/2) \in S_6 = [0, 1] \times [2, 3]$, pues, siendo $\mathbf{f}_2(u, v) = (1, u - 1, v)$, $\mathbf{f}_6(u, v) = (u, v - 2, 1)$, se tiene $\mathbf{f}(3/2, 1) = \mathbf{f}_2(3/2, 1) = (1, 1/2, 1) = \mathbf{f}_6(1, 5/2) = \mathbf{f}(1, 5/2)$.

En el punto $\mathbf{p}$ el vector normal a K es

$$N_{\mathbf{f}}(\mathbf{p}) = N_{\mathbf{f}_2}(\mathbf{p}) = \frac{\partial \mathbf{f}_2}{\partial u}(\mathbf{p}) \times \frac{\partial \mathbf{f}_2}{\partial v}(\mathbf{p}) = (0, 1, 0) \times (0, 0, 1) = (1, 0, 0)$$

y en el punto $\mathbf{p}'$ es

$$N_{\mathbf{f}}(\mathbf{p}') = N_{\mathbf{f}_6}(\mathbf{p}') = \frac{\partial \mathbf{f}_6}{\partial u}(\mathbf{p}') \times \frac{\partial \mathbf{f}_6}{\partial v}(\mathbf{p}') = (1, 0, 0) \times (0, 1, 0) = (0, 0, 1)$$

Por supuesto que los vectores $N_{\mathbf{f}}(\mathbf{p}) = (1, 0, 0)$ y $N_{\mathbf{f}}(\mathbf{p}') = (0, 0, 1)$ son linealmente independientes. Esto nos impide tener un vector normal bien definido para K en $\mathbf{q}$, por lo que no podemos hablar de un plano tangente al cubo en ese punto. Esto, además, es lo que la intuición nos dice que debería suceder, pues $\mathbf{q}$ está en una arista del cubo, donde la superficie deja de ser "suave" como para poderle asociar un plano tangente. ∎

Ejemplo 8. Consideremos la superficie K del ejemplo 2, parametrizada por $\mathbf{f}: S = [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (r \cos u \operatorname{sen} v, r \operatorname{sen} u \operatorname{sen} v, r \cos v)$. Los puntos $\mathbf{p} = (0, \pi/4)$ y $\mathbf{p}' = (2\pi, \pi/4)$ en S tienen la misma imagen $\mathbf{q} = \mathbf{f}(0, \pi/4) = (r/\sqrt{2}, 0, r/\sqrt{2}) = \mathbf{f}(2\pi, \pi/4)$.

Se tiene

$$\frac{\partial \mathbf{f}}{\partial u} = (-r \operatorname{sen} u \operatorname{sen} v, r \cos u \operatorname{sen} v, 0)$$
$$\frac{\partial \mathbf{f}}{\partial v} = (r \cos u \cos v, r \operatorname{sen} u \cos v, -r \operatorname{sen} v)$$

por lo que el vector normal es

$$N_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -r \operatorname{sen} u \operatorname{sen} v & r \cos u \operatorname{sen} v & 0 \\ r \cos u \cos v & r \operatorname{sen} u \cos v & -r \operatorname{sen} v \end{bmatrix}$$
$$= (-r^2 \cos u \operatorname{sen}^2 v, -r^2 \operatorname{sen} u \operatorname{sen}^2 v, -r^2 \operatorname{sen} v \cos v)$$

En el punto $\mathbf{p}$ se tiene

$$N_{\mathbf{f}}(\mathbf{p}) = N_{\mathbf{f}}(0, \pi/4) = \left(-\frac{r^2}{2}, 0, -\frac{r^2}{2}\right)$$

y en el punto $\mathbf{p}'$ se tiene

$$N_{\mathbf{f}}(\mathbf{p}') = N_{\mathbf{f}}(2\pi, \pi/4) = \left(-\frac{r^2}{2}, 0, -\frac{r^2}{2}\right) = N_{\mathbf{f}}(\mathbf{p})$$

Así pues, en el punto $\mathbf{q} = (r/\sqrt{2}, 0, r/\sqrt{2})$, imagen bajo $\mathbf{f}$ de los dos puntos $\mathbf{p}$ y $\mathbf{p}'$ de S, es posible trazar un plano tangente a la esfera (como era de esperarse). Este es

$$\left(-\frac{r^2}{2}, 0, -\frac{r^2}{2}\right) \cdot \left((x, y, z) - \left(\frac{r}{\sqrt{2}}, 0, \frac{r}{\sqrt{2}}\right)\right) = 0$$

o sea

$$x + z = \sqrt{2} r$$ ∎

Ejemplo 9. Sea $\mathbf{f}: S = [0, 2\pi] \times [0, h] \to \mathbb{R}^3$, $(h > 1)$,

$$\mathbf{f}(u, v) = (v \cos u, v \operatorname{sen} u, v)$$

La superficie $K = \mathbf{f}(S)$ es el cono del ejemplo 3. Tomemos el punto $\mathbf{q} = (1, 0, 1) \in K$, imagen de $\mathbf{p} = (0, 1)$ y $\mathbf{p}' = (2\pi, 1)$, puntos de la frontera de S. El vector normal $N_{\mathbf{f}}(u, v)$ es

$$N_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = (-v \operatorname{sen} u, v \cos u, 0) \times (\cos u, \operatorname{sen} u, 1)$$
$$= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -v \operatorname{sen} u & v \cos u & 0 \\ \cos u & \operatorname{sen} u & 1 \end{bmatrix} = (v \cos u, v \operatorname{sen} u, -v)$$

Se ve entonces que

$$N_{\mathbf{f}}(\mathbf{p}) = N_{\mathbf{f}}(0, 1) = (1, 0, -1) = N_{\mathbf{f}}(2\pi, 1) = N_{\mathbf{f}}(\mathbf{p}')$$

de modo que podemos trazar un plano tangente al cono en el punto $\mathbf{q}$, el cual es $z = x$. Por otra parte, en el punto $\mathbf{q} = (0, 0, 0)$, imagen de todos los puntos $(u, 0)$ de la frontera de S, los vectores normales son todos iguales a

$$N_{\mathbf{f}}(u, 0) = (0, 0, 0)$$

de tal modo que en el punto $\mathbf{q}$ no es posible tener un plano tangente al cono. Esto, por supuesto, era de esperarse, pues $\mathbf{q}$ es el vértice del cono. ∎

Ejercicios (Capítulo 8, Sección 4)

1. Sea $K_1 = \{(x, y, z) | z = x^2 + y^2 - 1, z \leq 0\}$, $K_2 = \{(x, y, z) | x^2 + y^2 + z^2 = 1, z \geq 0\}$. Demuestre que $K = K_1 \cup K_2$ es una superficie, dando una parametrización $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de ella. Describa geométricamente a K. Determine la ecuación del plano tangente a K en los puntos $\mathbf{q}_1 = (0, 0, -1)$ y $\mathbf{q}_2 = (0, 0, 1)$. ¿Es posible determinar un plano tangente a K en el punto (1,0,0)? Explique.

2. Sea $K_1 = \{(x, y, z) | z = e^{-(x^2+y^2)}, x^2 + y^2 \leq 1\}$, $K_2 = \{(x, y, z) | x^2 + y^2 = 1, -1 \leq z \leq e^{-1}\}$. Demuestre que $K = K_1 \cup K_2$ es una superficie, dando una parametrización $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de ella. Describa K en forma geométrica. Determine la ecuación del plano tangente a K en los puntos $\mathbf{q}_1 = (0, 0, 1)$ y $\mathbf{q}_2 = (\sqrt{2}/2, \sqrt{2}/2, 0)$. ¿Es posible determinar un plano tangente a K en el punto $(1, 0, e^{-1})$? Explique.

3. Sea $K_1 = \{(x, y, z) | x^2 + y^2 = 4, -2 \leq z \leq 0\}$, $K_2 = \{(x, y, z) | z = 4 - x^2 - y^2, 0 \leq z \leq 3\}$, $K_3 = \{(x, y, z) | x^2 + y^2 + (z - 3)^2 = 1, z \geq 3\}$. Demuestre que $K = K_1 \cup K_2 \cup K_3$ es una superficie, dando una parametrización $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de ella. Describa geométricamente K. ¿En qué puntos es posible trazar planos tangentes a K?

4. Sea K_1 la esfera con centro en el origen y radio 1, y sea K_2 la esfera con centro en el punto $(0, 1, 0)$ y radio 1. ¿Es $K = K_1 \cup K_2$ una superficie? Explique.

5. Sea K_1 la esfera con centro en el origen y radio 1, y sea K_2 la esfera con centro en el punto $(0, 2, 0)$ y radio 1. ¿Es $K = K_1 \cup K_2$ una superficie? Explique.

6. Sea $K = \{(x, y, z) \mid |x| + |y| + |z| = 1\}$. Demuestre que K es una superficie, dando una parametrización $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de ella. Describa en forma geométrica K. ¿En qué puntos es posible trazar planos tangentes a K? ¿Cuáles son estos? Explique.

7. Sea $K = \{(x, y, z) \mid z = |1 - x^2 - y^2|, x^2 + y^2 \leq 4\}$. Demuestre que K es una superficie, dando una parametrización $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de ella. Describa K en forma geométrica. ¿En qué puntos es posible trazar planos tangentes a K?

8. Sean K_1 y K_2 superficies simples, parametrizadas por $\mathbf{f}_1: S_1 \subset \mathbb{R}^2 \to \mathbb{R}^3$, y $\mathbf{f}_2: S_2 \subset \mathbb{R}^2 \to \mathbb{R}^3$, respectivamente. Suponga que $S_1 \cap S_2 = \varnothing$ y $K_1 \cap K_2 = \varnothing$. ¿Es $K = K_1 \cup K_2$ una superficie? Explique.

8.5 Orientación de superficies

En esta sección abordaremos el problema de asignar una orientación a una superficie. Así como en el caso de las curvas se tuvo que establecer el concepto de "curva positivamente orientada", pues éste apareció en el teorema de Green, ahora debemos tener un concepto de "superficie orientada", pues éste aparecerá en el teorema de Stokes, que se estudiará en el próximo capítulo.

En realidad lo que haremos en esta sección será simplemente dar la idea general de lo que en matemáticas se entiende por superficie orientable (y orientada): el hecho es que "la mayoría" de las superficies con las que nos podemos encontrar, en un curso de cálculo, tienen esta propiedad y no hay necesidad de entrar en muchos detalles técnicos para trabajar con ella cuando sea necesario (en general, la idea matemática de orientabilidad de superficies es fácil de entender intuitivamente, pero muy delicada cuando se quiere entrar en los detalles que la sostienen).

De manera informal, una superficie K en $\mathbb{R}^3$ se dirá "orientable" si es posible decidir sin ambigüedad cuál es cada uno de los lados de la superficie, por ejemplo, pintando cada uno de ellos con dos colores distintos. Si la superficie es simple, su frontera separará cada uno de los lados de la superficie. Si se trata de una superficie más general, por ejemplo un esfera, los lados de esta se identifican con su "interior" y su "exterior".

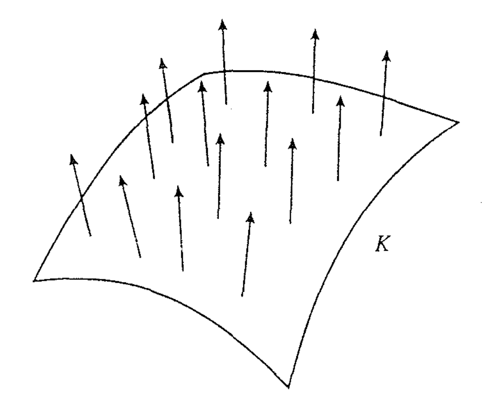


Figura 1. Los vectores normales a una superficie K

La herramienta con la que se puede hacer precisa esta idea general de orientabilidad es por medio de vectores normales a la superficie. La misión de ellos es que *apunten* en la dirección de *uno de los lados de la superficie*, distinguiéndolo de los demás lados de ella (figura 1).

Una función $\mathbf{N}: K \to \mathbb{R}^3$, definida en los puntos $\mathbf{q}$ de la superficie K, tal que a cada $\mathbf{q} \in K$ le asocia un vector $\mathbf{N}(\mathbf{q}) \in \mathbb{R}^3$, no nulo, ortogonal a K, se dice ser un *campo de vectores normales* a K (sin pérdida de generalidad podemos suponer que cada vector $\mathbf{N}(\mathbf{q})$ tiene norma 1).

Decir que una superficie K es orientable, significa que podemos tener un campo de vectores normales a K, $\mathbf{N}: K \to \mathbb{R}^3$, que no cambia repentinamente de un punto a otro de la superficie, es decir, que si un punto $\mathbf{q} \in K$ "está en un lado de K, sus vecinos sigan estando del mismo lado". En palabras más precisas: se requiere que este campo $\mathbf{N}$ sea *continuo* en K.

El primer ejemplo de superficie orientable es el caso de una superficie *simple* $K = \mathbf{f}(S)$, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. En efecto, hemos visto que en cada punto de K es posible determinar el vector no nulo

$$N_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}$$

el cual es un vector normal a la superficie. Entonces, si definimos $\mathbf{N}: K \to \mathbb{R}^3$ como

$$\mathbf{N}(x, y, z) = \mathbf{N}(\mathbf{f}(u, v)) = \frac{N_{\mathbf{f}}(u, v)}{\|N_{\mathbf{f}}(u, v)\|} = \frac{\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}}{\left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\|}$$

éste será un campo de vectores normales a K, el cual es continuo (pues siendo $\mathbf{f}$ de clase $\mathscr{C}^1$, sus derivadas parciales que aparecen en la última expresión son funciones continuas).

Siendo K una superficie orientable, el campo de vectores normales $\mathbf{N}: K \to \mathbb{R}^3$, le asigna una *orientación*. Así, la superficie K junto con el campo $\mathbf{N}$ determinan una *superficie orientada* (que en ocasiones se suele denotar como la pareja $(K, \mathbf{N})$).

Por ejemplo, si K es una superficie simple, podemos hacer ésta superficie orientada por medio del campo $\mathbf{N}: K \to \mathbb{R}^3$ definido anteriormente. Es decir, $(K, \mathbf{N})$ es una superficie orientada, donde $\mathbf{N}: K \to \mathbb{R}^3$ es

$$\mathbf{N}(x, y, z) = \frac{\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}}{\left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\|}$$

Queda, a este respecto, un detalle por aclarar. La orientación dada por el campo $\mathbf{N}$ está escrita en términos de una parametrización concreta de la superficie simple K. Así, esta orientación está comprometida con la parametrización $\mathbf{f}$ de K. Es natural preguntarse si un cambio de parametrización de K puede producir un cambio en la orientación dada por el campo $\mathbf{N}$ vía la parametrización $\mathbf{f}$. Veremos que, de hecho, esto ocurre. Primeramente observe que en cada punto $\mathbf{q} \in K$, un vector normal al plano tangente a K en $\mathbf{q}$, tiene solamente dos posibles direcciones, digamos: 1) la de $N_{\mathbf{f}}(u, v)$, ó 2) la de $-N_{\mathbf{f}}(u, v)$. Cada una de estas determina una orientación específica en la superficie K. Sea entonces $\mathbf{g} = \mathbf{f} \circ \boldsymbol{\varphi}: S' \subset \mathbb{R}^2 \to \mathbb{R}^3$ una reparametrización de K, donde $\boldsymbol{\varphi}: S' \to S$, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$, es una biyección de clase $\mathscr{C}^1$ cuyo determinante jacobiano $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)}$ es no nulo para todo (s, t) en S'. Con esta parametrización de K se consigue el campo de vectores normales a K, $\tilde{\mathbf{N}}: K \to \mathbb{R}^3$, a saber

$$\tilde{\mathbf{N}}(x, y, z) = \mathbf{N}_{\mathbf{g}}(s, t) = \frac{\dfrac{\partial \mathbf{g}}{\partial s} \times \dfrac{\partial \mathbf{g}}{\partial t}}{\left\| \dfrac{\partial \mathbf{g}}{\partial s} \times \dfrac{\partial \mathbf{g}}{\partial t} \right\|}$$

Los vectores $\mathbf{N}_{\mathbf{f}}(u, v)$ y $\mathbf{N}_{\mathbf{g}}(s, t)$ están relacionados por

$$\mathbf{N}_{\mathbf{g}}(s, t) = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} \mathbf{N}_{\mathbf{f}}(u, v)$$

(ver fórmula (Â) en la sección 2). Puesto que el jacobiano $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)}$ es *nunca nulo* en S', éste será siempre positivo o siempre negativo. En el primer caso los vectores $\mathbf{N}_{\mathbf{f}}(u, v)$ y $\mathbf{N}_{\mathbf{g}}(s, t)$ estarán en la misma dirección, y diremos entonces que **g** *es una reparametrización de K que conserva* su *orientación*. En el segundo caso, si los vectores $\mathbf{N}_{\mathbf{g}}(s, t)$ están en la dirección de $-\mathbf{N}_{\mathbf{f}}(u, v)$, diremos que **g** *es una reparametrización de K que invierte* su *orientación* (se entiende que estamos pensando en K como una superficie orientada, con la orientación dada por el campo $\mathbf{N}(x, y, z) = \mathbf{N}_{\mathbf{f}}(u, v)$, vía la parametrización **f**).

Ejemplo 1. Consideremos la superficie parabólica $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: S = \{(u, v)|u^2 + v^2 \leq 10\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$. Esta es una superficie orientable y la parametrización **f** de K le proporciona una orientación dada por el campo de vectores $\mathbf{N}: K \to \mathbb{R}^3$, $\mathbf{N}(x, y, z) = \frac{\mathbf{N}_{\mathbf{f}}(u,v)}{\|\mathbf{N}_{\mathbf{f}}(u,v)\|}$, en la que $\mathbf{N}_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = (1, 0, 2u) \times (0, 1, 2v) = (-2u, -2v, 1)$. Observe que estos vectores están apuntando hacia el interior de K. En el ejemplo 5 de la sección 3 se consideró la reparametrización $\mathbf{g}: S' = \{(s, t)|s^2 + t^2 \leq 5\} \to \mathbb{R}^3$, $\mathbf{g}(s, t) = (s + t, s - t, 2s^2 + 2t^2)$. Esta es, de hecho, la composición de **f** con la función $\varphi: S' \to S$, $\varphi(s, t) = (\varphi_1(s, t), \varphi_2(s, t)) = (s + t, s - t)$. El jacobiano de esta función es $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)} = -2$. Entonces **g** es una reparametrización de K que invierte la orientación de ésta. Por ejemplo, en el punto $\mathbf{q} = (1, 1, 2) = \mathbf{f}(1, 1) = \mathbf{g}(1, 0) \in K$, se tiene $\mathbf{N}_{\mathbf{f}}(1, 1) = (-2, -2, 1)$, el cual es un vector que apunta al interior de K en **q**, en tanto que $\mathbf{N}_{\mathbf{g}}(1, 0) = (4, 4, -2) = -2(-2, -2, 1) = -2\mathbf{N}_{\mathbf{f}}(1, 1)$, el cual es un vector que apunta al exterior de K. ■

Ejemplo 2. Sea K la porción del plano $2x + 2y + z = 1$ que se encuentra en el primer octante. Esta es una superficie simple parametrizada por $\mathbf{f}: S = \{(u, v)|0 \leq u \leq 1/2, 0 \leq v \leq 1/2 - u\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, 1 - 2u - 2v)$. Sea $S' = \{(s, t)| -1/2 \leq s \leq 0, 0 \leq t \leq s + 1/2\}$, y sea $\varphi: S' \to S$, $\varphi(s, t) = (\varphi_1(s, t), \varphi_2(s, t)) = (t, -s)$. Se verifica fácilmente que φ es una biyección de clase $\mathscr{C}^1$ cuyo determinante jacobiano es $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)} = 1$. Entonces la función $\mathbf{g} = \mathbf{f} \circ \varphi: S' \to \mathbb{R}^3$, $\mathbf{g}(s, t) = (t, -s, 1 - 2t + 2s)$ es una reparametrización de K que preserva su orientación (la que K tiene de la parametrización **f**). ■

Veamos ahora que también las superficies que se obtienen como niveles constantes de funciones $h: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ de clase $\mathscr{C}^1$ (sin puntos críticos) son superficies orientables.

De hecho, si $K = \{(x, y, z)|h(x, y, z) = c\}$ es el nivel c de h, sabemos (desde el capítulo 2) que el vector $\operatorname{grad} h(x, y, z)$ es normal a K en el punto (x, y, z). Así si definimos $\mathbf{N}: K \to \mathbb{R}^3$ como

$$\mathbf{N}(x, y, z) = \frac{\operatorname{grad} h(x, y, z)}{\|\operatorname{grad} h(x, y, z)\|} = \frac{\left(\dfrac{\partial h}{\partial x}, \dfrac{\partial h}{\partial y}, \dfrac{\partial h}{\partial z}\right)}{\sqrt{\left(\dfrac{\partial h}{\partial x}\right)^2 + \left(\dfrac{\partial h}{\partial y}\right)^2 + \left(\dfrac{\partial h}{\partial z}\right)^2}}$$

se tiene un campo de vectores normales en K, el cual es continuo, pues h es de clase $\mathscr{C}^1$, y determina una orientación de K.

Este último ejemplo junto con el hecho de que las gráficas de funciones de clase $\mathscr{C}^1$ de dos variables son orientables, nos hacen ver que "muchas" superficies son orientables. De hecho, todas las superficies que han aparecido hasta el capítulo anterior son orientables. Trabajaremos con ellas hasta el final del libro. Quisiéramos, sin embargo, terminar esta sección dando un ejemplo (el más célebre) de una superficie que *no* es orientable.

En el ejemplo 5 de la sección 4 se estudió la superficie $K = \mathbf{f}(S)$ parametrizada por $\mathbf{f}: S = [0, 2\pi] \times [-1, 1] \to \mathbb{R}^3$

$$\mathbf{f}(u, v) = \left(\left(1 - v \operatorname{sen} \frac{u}{2}\right) \cos u, \left(1 - v \operatorname{sen} \frac{u}{2}\right) \operatorname{sen} u, v \cos \frac{u}{2}\right)$$

la cual es conocida como "banda de Möbius". Esta es una superficie en $\mathbb{R}^3$ que no es orientable. Es decir, no es posible decidir en ella cuál es cada uno de sus lados. De hecho, se puede pensar en esta superficie como si tuviera un solo lado; si se comenzara a pintar "uno de sus lados" como se muestra en la figura, podríamos pintar *toda* la superficie sin "cambiar de lado"

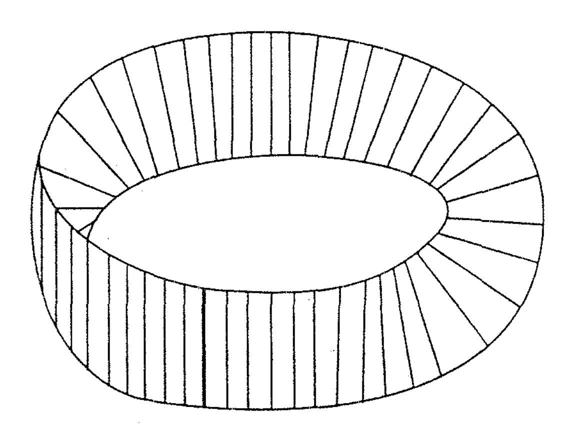


Figura 2. Se puede pintar toda la banda de Möbius sin cambiar de lado.

De otro modo: si intentamos sembrar en K un campo de vectores normales $\mathbf{N}: K \to \mathbb{R}^3$, por intuición se puede ver que éste no puede ser continuo. En efecto, se tienen puntos $\mathbf{q}$ y $\mathbf{q}'$ en K que están muy cerca uno de otro, pero que sus imágenes bajo $\mathbf{N}$, $\mathbf{N}(\mathbf{q})$ y $\mathbf{N}(\mathbf{q}')$, son vectores en $\mathbb{R}^3$ que están "muy lejos", pues apuntan de direcciones contrarias.

Recordemos que lo que hace la función $\mathbf{f}$ con el rectángulo $S = [0, 2\pi] \times [-1, 1]$ para formar (con sus imágenes) la banda de Möbius, es pegar las esquinas contrarias del rectángulo. Si desde S pensamos en un campo de vectores normales apuntando "hacia arriba", se ve que los puntos que están cerca de dos vértices opuestos de S quedarán, en $\mathbf{f}(S)$, "muy cerca", pero con sus vectores normales respectivos en direcciones contrarias.

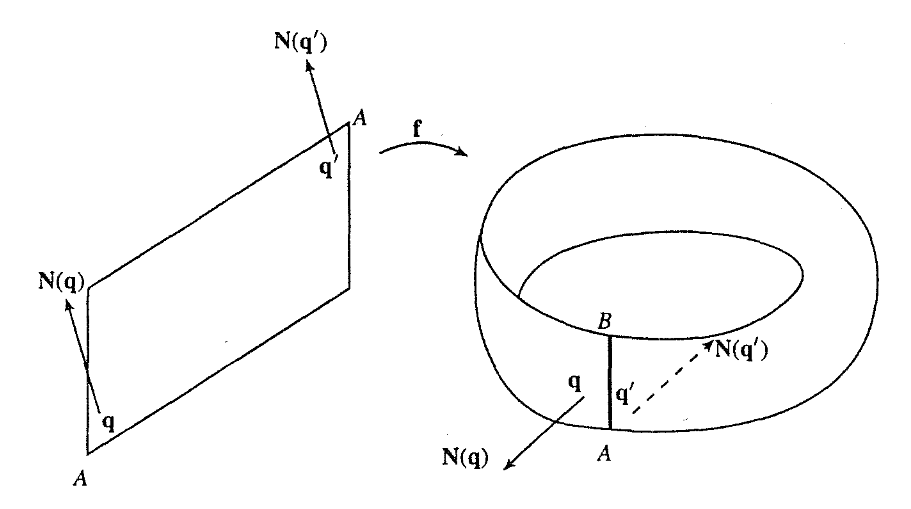


Figura 3. Los vectores normales en la banda de Möbius.

Para ver este hecho de modo más concreto , consideremos el campo de vectores normales correspondientes a los vectores $\mathbf{N_f} = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}$. Se tiene

$$\frac{\partial \mathbf{f}}{\partial u} = \left(-\frac{v}{2}\cos\frac{u}{2}\cos u - \left(1 - v\operatorname{sen}\frac{u}{2}\right)\operatorname{sen} u, \right.$$

$$\left. -\frac{v}{2}\cos\frac{u}{2}\operatorname{sen} u + \left(1 - v\operatorname{sen}\frac{u}{2}\right)\cos u, -\frac{v}{2}\operatorname{sen}\frac{u}{2}\right)$$

$$\frac{\partial \mathbf{f}}{\partial v} = \left(-\operatorname{sen}\frac{u}{2}\cos u, -\operatorname{sen}\frac{u}{2}\operatorname{sen} u, \cos\frac{u}{2}\right)$$

$$\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = \left(-\frac{v}{2}\operatorname{sen} u + \cos u \cos\frac{u}{2} - \frac{v}{2}\cos u \operatorname{sen} u, \right.$$

$$\left. \frac{v}{2}\cos u + \operatorname{sen} u \cos\frac{u}{2} - \frac{v}{2}\operatorname{sen}^2 u, \left(1 - v\operatorname{sen}\frac{u}{2}\right)\operatorname{sen}\frac{u}{2}\right)$$

los puntos

$$\mathbf{q} = \mathbf{f}(0.01, -0.99) = (1.0049, 0.01, -0.99)$$

y

$$\mathbf{q}' = \mathbf{f}(6.2732, 0.99) = (0.995, -0.01, -0.99)$$

son puntos de la banda de Möbius que están muy cerca el uno del otro (una estimación de su cercanía es $\|\mathbf{q} - \mathbf{q}'\| = 0.0223$). Sus correspondientes vectores normales $\mathbf{N}(\mathbf{q})$ y $\mathbf{N}(\mathbf{q}')$ son

$$\mathbf{N}(\mathbf{q}) = \frac{\mathbf{N_f}(\mathbf{q})}{\|\mathbf{N_f}(\mathbf{q})\|} = \frac{(1.0098, -0.485, 0.005)}{\|(1.0098, -0.485, 0.005)\|} = (0.9014, -0.4329, 0.0045)$$

$$\mathbf{N}(\mathbf{q}') = \frac{\mathbf{N_f}(\mathbf{q}')}{\|\mathbf{N_f}(\mathbf{q}')\|} = \frac{(0.99, -0.5049, 0.005)}{\|(0.99, -0.5049, 0.005)\|} = (-0.8908, 0.4543, 0.0045)$$

los cuales están muy lejos uno del otro (una estimación de su lejanía es $\|\mathbf{N}(\mathbf{q}) - \mathbf{N}(\mathbf{q}')\| = 1.9998$). Así pues, el campo $\mathbf{N}: K \longrightarrow \mathbb{R}^3$ no puede ser un campo continuo.

Ejercicios (Capítulo 8, Sección 5)

En los ejercicios 1–10 se da una superficie orientable K. Obtenga en cada caso una parametrización $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de K, que produzca la orientación indicada. (Es decir, de modo que el campo de vectores $\mathbf{N}(x, y, z) = \mathbf{N}(\mathbf{f}(u, v)) = \frac{\mathbf{N}_f(u,v)}{\|\mathbf{N}_f(u,v)\|} = \frac{\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}}{\|\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\|}$ produzca la orientación indicada).

1. $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1\}$, con sus vectores normales apuntando hacia afuera de la esfera K.

2. $K = \{(x, y, z)|x^2 + y^2 = 1, -1 \leq z \leq 1\}$, con sus vectores normales apuntando hacia adentro del cilindro K.

3. $K = \{(x, y, z)|z = 1 - x^2 - y^2, z \geq 0\}$, con sus vectores normales apuntando hacia adentro de la porción de paraboloide K (hacia donde está el origen de coordenadas).

4. $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, z \geq 0\}$, con sus vectores normales apuntando hacia adentro del casquete esférico K.

5. $K = \{(x, y, z)|x^2 + 3y^2 + 4z^2 = 1, z \leq 0\}$, con sus vectores normales apuntando hacia adentro del casquete elíptico K.

6. $K = \{(x, y, z) \mid |x| + |y| + |z| = 1\}$, con sus vectores normales apuntando hacia afuera de K.

7. $K = \{(x, y, z)|0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}$, con sus vectores normales apuntando hacia adentro de K.

8. $K = \{(x, y, z)|x + y + z = 1, x \geq 0, y \geq 0, z \geq 0\}$, con sus vectores normales apuntando hacia "atrás" del plano K (hacia el lado en que se encuentra el origen de coordenadas).

9. $K = \{(x, y, z)|z = e^{-(x^2+y^2)}, x^2 + y^2 \leq 1\}$, con sus vectores normales apuntando hacia "arriba" de K (hacia la parte positiva del eje z).

10. $K = \{(x, y, z)|z = x^2 + y^2 - 4, z \leq 0\} \cup \{(x, y, z)|z = 4 - x^2 - y^2, z \geq 0\}$, con sus vectores normales apuntando hacia afuera de K.

8.6 Área de una superficie

En el capítulo 6 se ha considerado el problema de calcular el área de ciertas regiones en el plano. Para ciertas superficies en el espacio, se estudió también el problema de determinar el "volumen bajo la superficie" o bien, el volumen encerrado por ella. En esta sección veremos cómo calcular el área de una superficie en $\mathbb{R}^3$.

Comencemos por considerar una superficie simple $K = \mathbf{f}(S)$, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, definida en la región simple del tipo I y II, S. Como primera aproximación al problema, pensemos que S es un rectángulo de dimensión $h_1 h_2$, cuyo vértice inferior izquierdo es el punto $\mathbf{p}_1 = (a, b)$, como se muestra en la figura 1.

La función $\mathbf{f}$ llevará el rectángulo S al espacio $\mathbb{R}^3$ y formará la superficie simple $K = \mathbf{f}(S)$. consideremos los caminos

$$\boldsymbol{\lambda}: [a, a + h_1] \to \mathbb{R}^3 \qquad \boldsymbol{\mu}: [b, b + h_2] \to \mathbb{R}^3$$
$$\boldsymbol{\lambda}(t) = \mathbf{f}(t, b) \qquad \boldsymbol{\mu}(t) = \mathbf{f}(a, t)$$

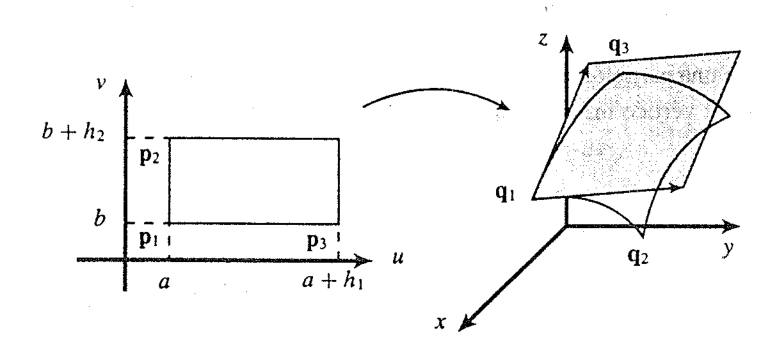


Figura 1. El rectángulo S y la función $\mathbf{f}$.

Las curvas descritas por estos caminos, imágenes de dos de los lados, $\overline{\mathbf{p}_1\mathbf{p}_2}$ y $\overline{\mathbf{p}_1\mathbf{p}_3}$, del rectángulo S respectivamente, corresponden a los "lados" $\mathbf{q}_1\mathbf{q}_2$ y $\mathbf{q}_1\mathbf{q}_3$ de la superficie simple K. Siendo $\mathbf{f}$ de clase $\mathscr{C}^1$, los caminos $\boldsymbol{\lambda}$ y $\boldsymbol{\mu}$ lo son también, de modo que podemos calcular la longitud como

$$|\mathbf{q}_1\mathbf{q}_2| = L_{\boldsymbol{\lambda}} = \int_a^{a+h_1} \|\boldsymbol{\lambda}'(t)\|\, dt \qquad |\mathbf{q}_1\mathbf{q}_3| = L_{\boldsymbol{\mu}} = \int_a^{a+h_2} \|\boldsymbol{\mu}'(t)\|\, dt$$

Según el teorema del valor medio para integrales, la primera de las dos integrales anteriores es igual a $\|\boldsymbol{\lambda}'(c)\|\, h_1$, en que c es un número entre entre a y $a+h_1$. Si el rectángulo S es suficientemente pequeño, podemos tener a c como si fuera a y obtener la expresión aproximada de la longitud de $\boldsymbol{\lambda}$ como

$$|\mathbf{q}_1\mathbf{q}_2| = L_{\boldsymbol{\lambda}} \approx \|\boldsymbol{\lambda}'(a)\| h_1$$

De igual modo se obtiene que

$$|\mathbf{q}_1\mathbf{q}_3| = L_{\boldsymbol{\mu}} \approx \|\boldsymbol{\mu}'(b)\| h_2$$

las cuales son expresiones que dan las normas de los vectores en $\mathbb{R}^3$

$$\mathbf{v}_1 = h_1 \frac{\partial \mathbf{f}}{\partial u}(a,b) \qquad \mathbf{v}_2 = h_2 \frac{\partial \mathbf{f}}{\partial v}(a,b)$$

Una medida aproximada del área de K, si el rectángulo S es suficientemente pequeño y $\mathbf{f}$ no lo deforma demasiado, es la norma del producto cruz de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$, la cual nos da el área del paralelogramo que estos vectores generan. Así

$$\begin{aligned}\text{Área de } K \approx \|\mathbf{v}_1 \times \mathbf{v}_2\| &= h_1 h_2 \left\| \frac{\partial \mathbf{f}}{\partial u}(a,b) \times \frac{\partial \mathbf{f}}{\partial v}(a,b) \right\| \\ &= \|\mathbf{N}_{\mathbf{f}}(a,b)\| h_1 h_2 = \|\mathbf{N}_{\mathbf{f}}(a,b)\| \text{ (área de } S)\end{aligned}$$

(un argumento análogo se usó en la sección 4 del capítulo 6 al dar argumentos geométricos de la fórmula de cambio de variables en integrales dobles).

Este preámbulo hecho con un rectángulo S, en el que se ha obtenido solamente una expresión aproximada del área de K, se usará ahora para establecer el caso general: se trata de dividir la región S en "rectangulitos", aplicar las ideas anteriores en cada uno de ellos, sumar y pasar al límite.

Sea entonces S una región del tipo I y II y sea $K = \mathbf{f}(S)$ una superficie simple parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Tomemos una partición de S en rectángulos R_{ij}^S, la cual a su vez produce una partición en "rectángulos deformados por la función $\mathbf{f}$" en la superficie K, digamos R_{ij}^K. Sea $\mathbf{p}_{ij}$ el vértice inferior izquierdo del rectángulo R_{ij}^S. Aplicando el argumento previo, podemos decir que

$$\boxed{\text{Área de } R_{\text{ij}}^k \approx \|\mathbf{N_f}(\mathbf{p}_{ij})\|(\text{ área de } R_{ij}^s)} \qquad E$$

de modo que

$$\text{área de } K \approx \sum_{i,j} \text{área de } R_{ij}^K = \sum_{i,j} \|\mathbf{N_f}(\mathbf{p}_{ij})\|(\text{area de } R_{ij}^s)$$

Nótese que la última sumatoria puede ser identificada como una suma de Riemann de la función $h(u, v) = \|\mathbf{N_f}(u, v)\|$ sobre la región $S \subset \mathbb{R}^2$. Haciendo que el número de rectángulos R_{ij}^S crezca (y que las dimensiones de los rectángulos tiendan a cero), la suma de Riemann anterior se convertirá en la integral doble de la función h sobre la región S. Así, resulta natural aceptar la definición siguiente.

Definición. Sea $K = \mathbf{f}(S)$ una superficie simple en $\mathbb{R}^3$, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. El área de la superficie K se define como la integral doble

$$\text{Área de} K = \iint_S \|\mathbf{N_f}(u, v)\|\, dudv$$

en la que $\mathbf{N_f} = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}$. ∎

Ejemplo 1. Tomemos la superficie parabólica $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 1\}$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$. Se tiene

$$\begin{aligned} \mathbf{N_f}(u, v) &= \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = (1, 0, 2u) \times (0, 1, 2v) \\ &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 2u \\ 0 & 1 & 2v \end{bmatrix} = (-2u, -2v, 1) \end{aligned}$$

de modo que $\|\mathbf{N_f}(u, v)\| = \sqrt{4u^2 + 4v^2 + 1}$, y así, el área de la superficie es

$$\begin{aligned} \text{Área de } K &= \iint_S \|\mathbf{N_f}(u, v)\|\, dudv = \iint_S \sqrt{4u^2 + 4v^2 + 1}\, dudv \\ &\overset{\text{polares}}{\longrightarrow} = \int_0^{2\pi} \int_0^1 \sqrt{1 + 4r^2} r\, dr d\theta = \frac{\pi}{6}(5^{3/2} - 1) \end{aligned}$$ ∎

Cuando la superficie simple K es la gráfica de una función $F: S \subset \mathbb{R}^2 \to \mathbb{R}$ de clase $\mathscr{C}^1$ (como en el ejemplo anterior), la fórmula de la definición dada arriba toma un aspecto especial. Se tiene en este caso que, siendo $\mathbf{f}(u, v) = (u, v, F(u, v))$ una parametrización de K,

$$\|\mathbf{N_f}(u, v)\| = \left\| - \left(\frac{\partial F}{\partial u}, -\frac{\partial F}{\partial v}, 1\right)\right\| = \sqrt{1 + \left(\frac{\partial F}{\partial u}\right)^2 + \left(\frac{\partial F}{\partial v}\right)^2}$$

(ver ejemplo 6 de la sección 3). Entonces el área de K, gráfica de $z = F(u, v)$ definida sobre la región S del tipo I y II se ve como

$$\text{Área de } K = \iint_S \sqrt{1 + \left(\frac{\partial F}{\partial u}\right)^2 + \left(\frac{\partial F}{\partial v}\right)^2}\, dudv$$

Ejemplo 2. Consideremos la superficie $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: S = [a, b] \times [c, d] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$ (ver ejemplo 1 de la sección 1). K es una porción de un plano en $\mathbb{R}^3$, gráfica de la función $F: S \subset \mathbb{R}^2 \to \mathbb{R}$, $F(u, v) = \alpha u + \beta v + \gamma$, sobre el rectángulo $S = [a, b] \times [c, d]$. El área de K es entonces

$$\begin{aligned}\text{Área de } K &= \int_S \sqrt{1 + \left(\frac{\partial F}{\partial u}\right)^2 + \left(\frac{\partial F}{\partial v}\right)^2}\, dudv \\ &= \int_a^b \int_c^d \sqrt{1 + \alpha^2 + \beta^2}\, dvdu = \sqrt{1 + \alpha^2 + \beta^2}(b - a)(d - c)\end{aligned}$$ ■

Ejemplo 3. Sea K la porción del casquete esférico $z = \sqrt{R_1^2 - x^2 - y^2}$ cortado por el cilindro circular recto $x^2 + y^2 = R_2^2$, $R_1 > R_2$. Esta es una superficie simple (¿por qué?; ver ejemplo 4 de la sección 1). Se trata entonces de la gráfica de la función $F(u, v) = \sqrt{R_1^2 - u^2 - v^2}$ sobre la región $S = \{(u, v) | u^2 + v^2 \leq R_2^2\}$. El área de esta superficie es

$$\begin{aligned}\text{Área de} K &= \iint_S \sqrt{1 + \left(\frac{\partial F}{\partial u}\right)^2 + \left(\frac{\partial F}{\partial v}\right)^2}\, dudv \\ &= \iint_S \sqrt{1 + \frac{u^2}{R_1^2 - u^2 - v^2} + \frac{v^2}{R_1^2 - u^2 - v^2}}\, dudv \\ &= \iint_S \frac{R_1}{\sqrt{R_1^2 - u^2 - v^2}}\, dudv \overset{\text{polares}}{=\!=} R_1 \int_0^{2\pi} \int_0^{R_2} \frac{r\, drd\theta}{\sqrt{R_1^2 - r^2}} \\ &= 2\pi R_1 \left(R_1 - \sqrt{R_1^2 - R_2^2}\right)\end{aligned}$$ ■

Ejemplo 4. Sea K la mitad superior de la esfera unitaria $x^2 + y^2 + z^2 = 1$ (ver ejemplo 4 de la sección 1). Una parametrización de K está dada por la proyección estereográfica $\mathbf{f}: S = \{(u, v) | u^2 + v^2 \leq 1\} \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = \dfrac{1}{1 + u^2 + v^2}(2u, 2v, 1 - u^2 - v^2)$. Los vectores $\dfrac{\partial \mathbf{f}}{\partial u}$ y $\dfrac{\partial \mathbf{f}}{\partial v}$ son

$$\begin{aligned}\frac{\partial \mathbf{f}}{\partial u} &= \frac{1}{(1 + u^2 + v^2)^2}(2(1 - u^2 + v^2), -4uv, -4u) \\ \frac{\partial \mathbf{f}}{\partial v} &= \frac{1}{(1 + u^2 + v^2)^2}(-4uv, 2(1 + u^2 - v^2), -4v)\end{aligned}$$

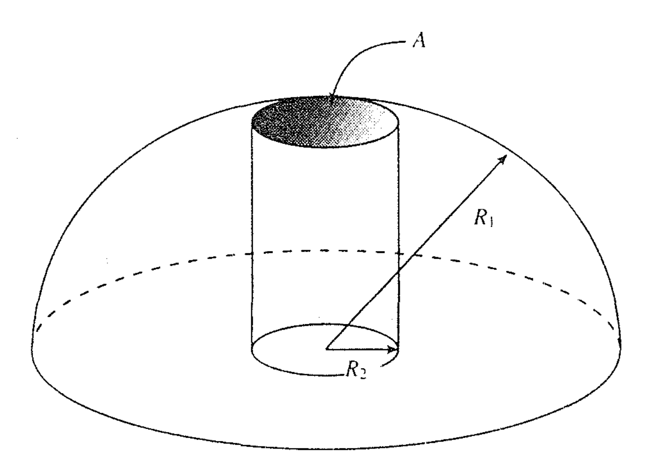


Figura 2. Área de la superficie del ejemplo 3.

cuyo producto cruz es

$$\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = \frac{1}{(1 + u^2 + v^2)^3}(8u, 8v, 4(1 - u^2 - v^2))$$

La norma de este vector es

$$\|\mathbf{N_f}(u, v)\| = \frac{4}{(1 + u^2 + v^2)^2}$$

Entonces el área de K es

$$\text{Área de } K = \iint_S \|\mathbf{N_f}(u, v)\|\, du dv = \iint_S \frac{4}{(1 + u^2 + v^2)^2}\, du dv$$

$$\overset{\text{polares}}{=\!=\!\longrightarrow} = 4 \int_0^{2\pi} \int_0^1 \frac{r\, dr d\theta}{(1 + r^2)^2} = 2\pi$$

Esta es el área de la mitad de la esfera unitaria (como ya sabíamos). Obsérvese que del resultado del ejemplo anterior también podemos obtener este valor con $R_1 = R_2 = 1$. Sin embargo, hasta este momento la teoría desarrollada no permite obtener el resultado de esta manera (¿por qué?). ■

Antes de continuar es necesario aclarar un detalle técnico de la definición de área de una supericie simple K establecida anteriormente: en ella aparece el área de K comprometida con una *parametrización concreta* $\mathbf{f}$ de la superficie. Es de esperarse que en realidad el área de K sea una propiedad de la superficie misma, y no dependa de la función que la parametriza. Esto, efectivamente es cierto, y ahora vamos a verificarlo. Sea entonces $\mathbf{g} = \mathbf{f} \circ \boldsymbol{\varphi}: S' \subset \mathbb{R}^2 \to \mathbb{R}^3$ una reparametrización de K, donde $\boldsymbol{\varphi}: S' \to S$, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$, es una biyección de clase $\mathscr{C}^1$ tal que $\det \boldsymbol{\varphi}'(s, t) = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} \neq 0$ $\forall (s, t) \in S'$. Usaremos el teorema de cambio de variables en integrales dobles: la idea es cambiar

las variables u y v de $\mathbf{N_f}(u, v)$ por las variables s y t del dominio de φ. Se tiene

$$\underset{\text{según } \mathbf{f}}{\text{Área de } K} = \iint_S \|\mathbf{N_f}(u, v)\|\, du\, dv = \xrightarrow[(u,v)=\varphi(s,t)]{\text{cambio de variables}}$$

$$= \iint_{S'} \|\mathbf{N_f}(\varphi(s, t))\| \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} \right| dsdt$$

Ahora bien, se había visto que los vectores $\mathbf{N_f}(u, v)$ y $\mathbf{N_g}(s, t)$ estaban relacionados por

$$\mathbf{N_g}(s, t) = \frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} \mathbf{N_f}(u, v)$$

de modo que la última integral es precisamente igual a

$$\iint_{S'} \|\mathbf{N_g}(s, t)\|\, ds\, dt = \underset{(\text{según } \mathbf{g})}{\text{Área de } K}$$

lo que muestra que el área de la superficie K es independiente de la parametrización que se tenga de ella.

Ejemplo 5. Retomemos el ejemplo 1. Es fácil verificar que una reparametrización de la superficie parabólica $K = \mathbf{f}(S)$ ahí considerada (parametrizada por $\mathbf{f}: S = \{(u, v) | u^2 + v^2 \leq 1\} \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$), es $\mathbf{g}: S' = \{(s, t) | s^2 + t^2 \leq 1/2\} \to \mathbb{R}^3$, $\mathbf{g}(s, t) = (s + t, s - t, 2s^2 + 2t^2)$, (ver ejemplo 5 de la sección 3). Con esta reparametrización tenemos

$$\mathbf{N_g}(s, t) = \frac{\partial \mathbf{g}}{\partial s} \times \frac{\partial \mathbf{g}}{\partial t} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 4s \\ 1 & -1 & 4t \end{bmatrix} = (4t + 4s, 4s - 4t, -2)$$

de modo que el área de K es

$$\begin{aligned} \text{Área de } K &= \iint_{S'} \|\mathbf{N_g}(s, t)\|\, ds\, dt \\ &= \iint_{S'} \sqrt{(4t + 4s)^2 + (4s - 4t)^2 + (-2)^2}\, ds\, dt \\ &= \iint_{S'} \sqrt{32t^2 + 32s^2 + 4}\, ds\, dt \\ &= \underset{\text{polares}}{\longrightarrow} = 2 \int_0^{2\pi} \int_0^{1/\sqrt{2}} \sqrt{1 + 8r^2}\, r\, dr\, d\theta = \frac{\pi}{6}(5^{3/2} - 1) \end{aligned}$$

resultado que coincide con el del ejemplo 1. ∎

Las ideas anteriores para el cálculo del área de una superficie simple funcionan igualmente bien para el cálculo del área de superficies seccionalmente simples: si S es la unión de las regiones del tipo I y II, $S_1, S_2, \ldots, S_m$ y $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ parametriza a la superficie $K = \mathbf{f}(S)$, entonces

$$\text{Área de } K = \iint_S \|\mathbf{N_f}(u, v)\|\, du\, dv = \sum_{i=1}^{m} \iint_{S_i} \|\mathbf{N_f}(u, v)\|\, du\, dv$$

Ejemplo 6. En el ejemplo 6 de la sección 1, se consideró el cilindro $K = \{(x, y, z)|x^2 + y^2 = 1, -1 \leq z \leq 1\}$, parametrizado por $\mathbf{f}: S = \{(u, v)|1 \leq u^2 + v^2 \leq 9\} \subset \mathbb{R}^2 \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = \left(\frac{u}{\sqrt{u^2+v^2}}, \frac{v}{\sqrt{u^2+v^2}}, \sqrt{u^2+v^2} - 2\right)$$

Se trata de una superficie seccionalmente simple, para la cual

$$\begin{aligned}
\mathbf{N_f}(u, v) &= \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \\
&= \left(\frac{v^2}{(u^2+v^2)^{3/2}}, \frac{-uv}{(u^2+v^2)^{3/2}}, \frac{u}{\sqrt{u^2+v^2}}\right) \times \left(\frac{-uv}{(u^2+v^2)^{3/2}}, \frac{u^2}{(u^2+v^2)^{3/2}}, \frac{v}{\sqrt{u^2+v^2}}\right) \\
&= \left(-\frac{u}{u^2+v^2}, -\frac{v}{u^2+v^2}, 0\right)
\end{aligned}$$

de modo que

$$\|\mathbf{N_f}(u, v)\| = \frac{1}{\sqrt{u^2+v^2}}$$

Para calcular la integral doble sobre S de esta función, donde S es el anillo circular $1 \leq u^2 + v^2 \leq 9$ (que es una región que se descompone como unión de regiones del tipo I y II), conviene realizar el cambio a coordenadas polares. Se tiene

$$\text{Área de } K = \iint_S \frac{1}{\sqrt{u^2+v^2}}\, dudv \overset{\text{polares}}{\longrightarrow} = \int_0^{2\pi} \int_1^3 dr d\theta = 4\pi$$ ■

En la discusión previa a la definición presentada para el área de una superficie simple $K = \mathbf{f}(S)$, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ (cuyas características eran: inyectividad, de clase $\mathscr{C}^1$ e independencia lineal de los vectores $\frac{\partial \mathbf{f}}{\partial u}$ y $\frac{\partial \mathbf{f}}{\partial v}$ en todos los puntos de S), se trató de que la fórmula dada en tal definición fuera "natural". Las características de la función $\mathbf{f}$ hacen que el integrando que aparece en la integral doble con la que se calcula el área de K, es decir, la función $h(u, v) = \|\mathbf{N_f}(u, v)\|$ sea continuo en S. Sin embargo, obsérvese que tal fórmula hace perfecto sentido si el integrando $\|\mathbf{N_f}(u, v)\|$ tiene un "mal comportamiento" (v.gr. es discontinuo) en un subconjunto $\tilde{S}$ de S que tenga medida cero (ver nota 2 en la sección 3 del capítulo 6). Así, podemos aceptar que la función $\mathbf{f}$ pierda alguna o algunas de sus propiedades que la caracterizan como parametrización de K, siempre y cuando esto ocurra en un subconjunto de S de medida cero. Esto nos permite establecer *la misma* definición de área (dada para superficies simples) para superficies más generales (como las estudiadas en la sección 4).

Ejemplo 7. En el ejemplo 2 de la sección 4 se vio que la función $\mathbf{f}: S = [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (r \cos u \operatorname{sen} v, r \operatorname{sen} u \operatorname{sen} v, r \cos v)$, es una parametrización de la esfera $x^2 + y^2 + z^2 = r^2$, (la cual falla de ser inyectiva en los puntos de la frontera de S). Podemos, sin embargo, usarla para calcular el área de la esfera como lo hemos venido haciendo en esta sección. Se tiene

$$\|\mathbf{N_f}(u, v)\| = r^2 \operatorname{sen} v$$

por lo que el área de $K = \mathbf{f}(S)$ es

$$\text{Área de } K = \iint_S \|\mathbf{N_f}(u, v)\|\, du\, dv = \int_0^{2\pi} \int_0^{\pi/2} r^2 \operatorname{sen} v\, du\, dv = 2\pi r^2 \qquad \blacksquare$$

Ejemplo 8. En el ejemplo 4 de la sección 4 se vio que la función $\mathbf{f}: S = [0, 2\pi] \times [-1, 1] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (\cos u, \operatorname{sen} u, v)$, es una parametrización del cilindro $K = \mathbf{f}(S) = \{(x, y, z) | x^2 + y^2 = 1, -1 \leq z \leq 1\}$. En este caso se tiene

$$\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\operatorname{sen} u & \cos u & 0 \\ 0 & 0 & 1 \end{bmatrix} = (\cos u, \operatorname{sen} u, 0)$$

de modo que $\mathbf{N_f}(u, v) = 1$, y entonces

$$\text{Área de } K = \iint_S \|\mathbf{N_f}(u, v)\|\, du\, dv = \int_0^{2\pi} \int_{-1}^{1} dv\, du = 4\pi \qquad \blacksquare$$

Ejemplo 9. (Área del toro). Consideremos el toro $K = \mathbf{f}(S)$, imagen de la función $\mathbf{f}: S = [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = ((R + r \cos v) \cos u, (R + r \cos v) \operatorname{sen} u, r \operatorname{sen} v)$$

(ver ejemplo 6 de la sección 4). Hallemos el área de esta superficie. Se tiene

$$\begin{aligned}
\frac{\partial \mathbf{f}}{\partial u} &= (-(R + r \cos v) \operatorname{sen} u, (R + r \cos v) \cos u, 0) \\
\frac{\partial \mathbf{f}}{\partial v} &= (-r \operatorname{sen} v \cos u, -r \operatorname{sen} v \operatorname{sen} u, r \cos v) \\
\mathbf{N_f}(u, v) &= \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \\
&= \big(r \cos u \cos v(R + r \cos v), -r \operatorname{sen} u \cos v(R + r \cos v), r \operatorname{sen} v(R + r \cos v)\big) \\
\|\mathbf{N_f}(u, v)\| &= r(R + r \cos v)
\end{aligned}$$

de modo que

$$\text{Área de } K = \iint_S \|\mathbf{N_f}(u, v)\|\, du\, dv = \int_0^{2\pi} \int_0^{2\pi} r(R + r \cos v)\, du\, dv = 4\pi^2 Rr \qquad \blacksquare$$

Ejemplo 10. (Área de superficies de revolución. Otro teorema de Pappus). Sea C una curva simple en el plano xy, digamos que en el semiplano superior $y \geq 0$. Al girar esta curva alrededor del eje x obtenemos una superficie de revolución cuya área se puede calcular como $2\pi d\ell$, donde d es la distancia del centro de masa de la curva al eje de rotación y ℓ es la longitud de C. Este resultado es uno de los llamados "Teoremas de Pappus" (ver también ejemplo 12 de la sección 5 del capítulo 6). Para ver su validez tomemos una curva simple de clase $\mathscr{C}^1$ en el plano yz, $y \geq 0$, imagen del camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (y(t), z(t))$, $y(t) \geq 0\ \forall t \in [a, b]$ y pongámosla a girar alrededor del eje z, como se muestra en la figura 3

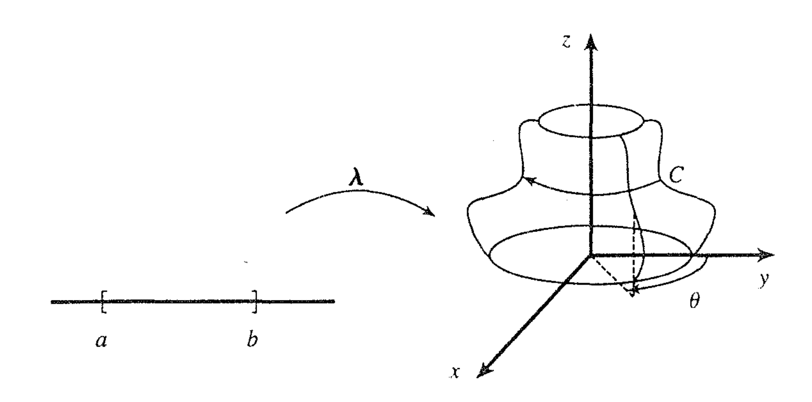


Figura 3. La superficie de revolución del ejemplo 10.

Nótese que una parametrización de la superficie de revolución así generada es $\mathbf{f}: [a, b] \times [0, 2\pi] \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(t, \theta) = (y(t) \operatorname{sen} \theta, y(t) \cos \theta, z(t))$. Se tiene

$$\frac{\partial \mathbf{f}}{\partial t} = (y'(t) \operatorname{sen} \theta, y'(t) \cos \theta, z'(t))$$

$$\frac{\partial \mathbf{f}}{\partial \theta} = (y(t) \cos \theta, -y(t) \operatorname{sen} \theta, 0)$$

$$\mathbf{N_f}(t, \theta) = \frac{\partial \mathbf{f}}{\partial t} \times \frac{\partial \mathbf{f}}{\partial \theta} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ y'(t) \operatorname{sen} \theta & y'(t) \cos \theta & z'(t) \\ y(t) \cos \theta & -y(t) \operatorname{sen} \theta & 0 \end{bmatrix}$$

$$= \big(y(t)z'(t) \operatorname{sen} \theta, y(t)z'(t) \cos \theta, -y(t)y'(t)\big)$$

$$\|\mathbf{N_f}(t, \theta)\| = y(t)\sqrt{(y'(t))^2 + (z'(t))^2} = y(t)\|\boldsymbol{\lambda}'(t)\|$$

entonces el área A de la superficie es

$$A = \iint_S y(t)\|\boldsymbol{\lambda}'(t)\|\, dt\, d\theta = \int_a^b \int_0^{2\pi} y(t)\|\boldsymbol{\lambda}'(t)\|\, d\theta\, dt$$

$$= 2\pi \int_a^b y(t)\|\boldsymbol{\lambda}'(t)\|\, dt$$

Nótese que la distancia d del centro de masa de la curva C, digamos que se localiza en el punto $(\bar{y}, \bar{z})$ del plano yz, al eje de rotación, es justamente la coordenada $\bar{y}$. Sabemos que ésta se calcula como

calcula como (ver sección 7 —subsección 2—, capítulo 7)

$$\bar{y} = \frac{\int_{\boldsymbol{\lambda}} y\rho(y, z)\, ds}{\int_{\boldsymbol{\lambda}} \rho(y, z)\, ds}$$

(donde las integrales están en relación con la longitud de arco). Siendo la curva C homogénea ($\rho(y, z) = \text{cte}$), podemos escribir esta expresión como

$$d = \frac{\int_{\boldsymbol{\lambda}} y\, ds}{\int_{\boldsymbol{\lambda}} ds} = \frac{\int_a^b y(t)\|\boldsymbol{\lambda}'(t)\|\, dt}{\int_a^b \|\boldsymbol{\lambda}'(t)\|\, dt}$$

de donde

$$\int_a^b y(t)\|\boldsymbol{\lambda}'(t)\|\, dt = d\int_a^b \|\boldsymbol{\lambda}'(t)\|\, dt = d\ell$$

y así, el área de la superficie de revolución considerada es

$$A = 2\pi d\ell$$

como queríamos demostrar.

Por ejemplo, sabemos que el centro de masa del semicírculo superior $y = \sqrt{r^2 - x^2}$ se encuentra en el punto $(0, 2r/\pi)$ —ver ejemplo 9, sección 7, capítulo 7—. Es claro que la superficie de revolución que este semicírculo genera es una esfera de radio r. Su área es

$$A = 2\pi d\ell = (2\pi)(2r/\pi)(\pi r) = 4\pi r^2$$

como ya sabíamos.

Nótese, por último, que el área del tubo calculada en el ejemplo 9, se obtiene de manera inmediata con el teorema de Pappus establecido en este ejercicio. En efecto, se trata de un círculo de radio r, cuyo centro gira en un círculo de radio R. El centro de masa del círculo que gira está en su centro (geométrico) y por lo tanto la distancia del centro de masa al eje de rotación es justamente R. Puesto que la longitud del círculo que gira es $2\pi r$, se tiene que

$$\text{Área del tubo } = 2\pi d\ell = 2\pi(R)(2\pi r) = 4\pi^2 Rr$$

como se obtuvo en el ejemplo 9. ■

Ejercicios (Capítulo 8, Sección 6)

1. Halle el área de la parte del plano $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, donde a, b, c son números positivos dados, que se se encuentra en el primer octante.

2. Halle el área de la parte de la esfera $x^2 + y^2 + z^2 = a^2$ que se encuentra dentro del cilindro $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, b \leq a, z \geq 0$.

3. Halle el área de la parte de la esfera $x^2 + y^2 + z^2 = a^2$ que se encuentra dentro del cilindro $x^2 + y^2 = ax$.

4. Halle el área de la parte del cono $z^2 = x^2 + y^2$, $z \geq 0$, que se encuentra dentro del cilindro $x^2 + y^2 = 2x$.

5. Halle el área de la superficie del cuerpo limitado por los cilindros $x^2 + y^2 = a^2$, $y^2 + z^2 = a^2$.

6. Halle el área de la parte del cono $z^2 = 3(x^2 + y^2)$, $z \geq 0$, que se encuentra por debajo del plano $x + y + z = 2a$, $a > 0$.

7. Halle el área de la parte de la esfera $x^2 + y^2 + z^2 = a^2$ en el primer octante, limitada por los cuatro planos $z = c$, $z = d$, $y = \alpha x$, $y = \beta x$, en donde $-a \leq c < d \leq a$, $0 < \alpha < \beta$.

8. Halle el área de la parte de la superficie $z = xy$ que se encuentra dentro del cilindro $x^2 + y^2 = a^2$.

9. (Teorema de Pitágoras generalizado a planos en $\mathbb{R}^3$). Considere la superficie simple $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: S = [a, b] \times [c, d] \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$. K es entonces la porción del plano en $\mathbb{R}^3$, $z = \alpha x + \beta y + \gamma$, en la región S de $\mathbb{R}^2$ (suponga que el plano no es paralelo a ninguno de los planos coordenados). Llame a los vértices del plano como sigue

$$P_1 = (a, c, \alpha a + \beta c + \gamma) \qquad P_2 = (b, c, \alpha b + \beta c + \gamma)$$
$$P_3 = (b, d, \alpha b + \beta d + \gamma) \qquad P_4 = (a, d, \alpha a + \beta d + \gamma)$$

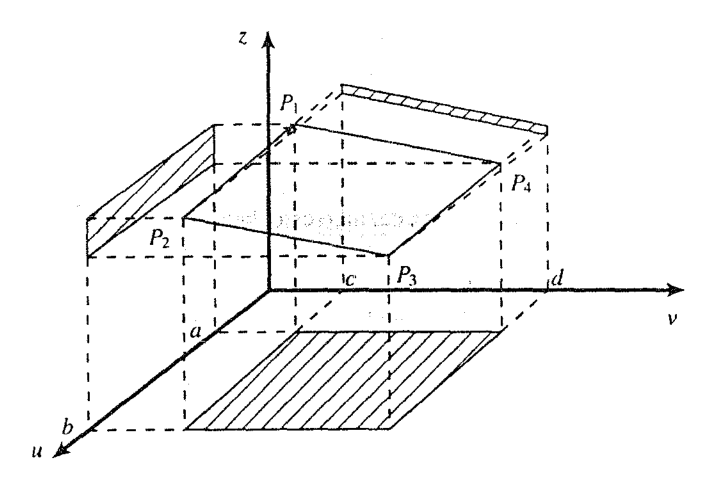


Figura 4. El plano $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$.

Denote por P_i^{π} a la proyección (ortogonal) del vértice P_i en el plano coordenado Π, $i = 1, 2, 3, 4$. Así, por ejemplo

$$P_1^{uv} = (a, c, 0) \qquad P_2^{uv} = (b, c, 0)$$
$$P_3^{uv} = (b, d, 0) \qquad P_4^{uv} = (a, d, 0)$$

a. Compruebe que $P_1^{uz} = (a, 0, \alpha a + \beta c + \gamma)$, $P_2^{uz} = (b, 0, \alpha b + \beta c + \gamma)$, $P_3^{uz} = (b, 0, \alpha b + \beta d + \gamma)$, $P_4^{uz} = (a, 0, \alpha a + \beta d + \gamma)$, $P_1^{vz} = (0, c, \alpha a + \beta c + \gamma)$, $P_2^{vz} = (0, c, \alpha b + \beta c + \gamma)$, $P_3^{vz} = (0, d, \alpha b + \beta d + \gamma)$, $P_4^{vz} = (0, d, \alpha a + \beta d + \gamma)$.

b. Demuestre que el área del paralelogramo cuyos vértices son las proyecciones P_1^{uz}, P_2^{uz}, P_3^{uz}, P_4^{uz} es igual a $\beta(b - a)(d - c)$.

c. Demuestre que el área del paralelogramo cuyos vértices son las proyecciones P_1^{vz}, P_2^{vz}, P_3^{vz}, P_4^{vz} es igual a $\alpha(b - a)(d - c)$.

d. En el ejemplo 2 de esta sección se obtuvo que el área de K es igual a $\sqrt{1 + \alpha^2 + \beta^2}(b - a)(d - c)$. Demuestre entonces que el cuadrado del área de K es igual a la suma de los cuadrados de las áreas de las 3 proyecciones de K en los planos coordenados uv, uzy vz. Esta es la versión en $\mathbb{R}^3$ del Teorema de Pitágoras.

10. Considere la superficie simple K parametrizada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Demuestre que el área de K puede ser calculada por la fórmula

$$\text{Área de } K = \iint_S \sqrt{EG - F^2}\, du dv$$

en donde

$$E = \left\| \frac{\partial \mathbf{f}}{\partial u} \right\|^2, \quad F = \frac{\partial \mathbf{f}}{\partial u} \cdot \frac{\partial \mathbf{f}}{\partial v}, \quad G = \left\| \frac{\partial \mathbf{f}}{\partial v} \right\|^2$$

11. Use el teorema de Pappus estudiado en el ejemplo 10 de esta sección para demostrar que el área (lateral) de un cono circular recto de radio de la base r y altura h es igual a $\pi r \sqrt{r^2 + h^2}$.

(*) 8.7 Tubos

En esta sección, que es de carácter opcional, estudiaremos algunos resultados interesantes sobre un tipo de objetos matemáticos llamados tubos. Con lo estudiado en la sección 8 del capítulo 5 (en la que se estudiaron las curvas paralelas) se tienen las bases necesarias para abordar el estudio de los tubos en $\mathbb{R}^2$, y lo que se ha discutido en este capítulo, junto con los resultados generales de curvas del capítulo 5, proporcionan los elementos para estudiar los tubos en $\mathbb{R}^3$. Estos son los dos únicos casos que consideraremos (en los cuales, además, se pueden "visualizar" los tubos).

8.7.1 Tubos en R^2

Sea $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^2$, una curva regular. En la sección 8 del capítulo 5 se estudió la curva $\boldsymbol{\beta}: [a, b] \to \mathbb{R}^2$ dada por

$$\boldsymbol{\beta}(t) = \boldsymbol{\alpha}(t) + r\mathbf{n}(t)$$

donde $\mathbf{n}(t)$ es el vector normal unitario que se obtiene al girar el vector tangente $\boldsymbol{\alpha}'(t)$ (normalizado) un ángulo de $\pi/2$ en dirección antihoraria. Esta curva $\boldsymbol{\beta}$ es la curva r-paralela a $\boldsymbol{\alpha}$, cuyo vector tangente $\boldsymbol{\beta}'(t)$ se puede calcular como

$$\boldsymbol{\beta}'(t) = (1 - rk_{\boldsymbol{\alpha}}(t))\boldsymbol{\alpha}'(t)$$

donde $k_\alpha(t)$ es la curvatura de α en t.

Para $r = \pm R$, con R suficientemente pequeño, obtenemos dos curvas paralelas a α, para las que α corre "por en medio de las dos". Al conjunto

$$T_{\mathbb{R}^2}(\alpha, R) = \{\beta([a, b]) | \beta \text{ es } r\text{-paralela a } \alpha, \ -R \leq r \leq R\}$$

le llamaremos *tubo* de centro en α y radio R

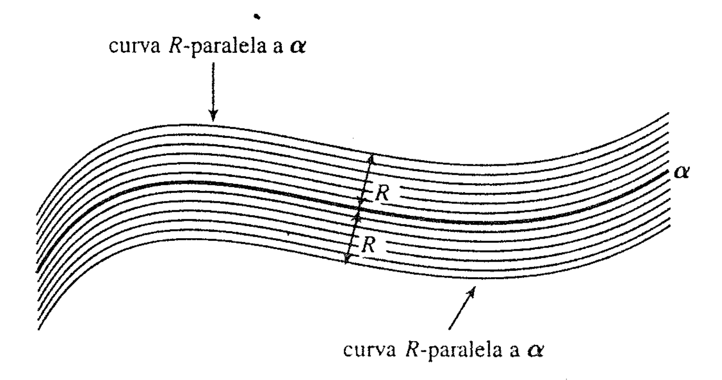


Figura 1. Tubo con centro en α y radio R

La longitud de cada curva β, r-paralela a α, que denotaremos por $L_r(\beta)$, con $-R \leq r \leq R$, se calcula como

$$L_r(\beta) = \int_a^b \|\beta'(t)\| \, dt = \int_a^b |1 - rk_\alpha(t)| \|\alpha'(t)\| \, dt$$

Como r es pequeño y $1 - rk_\alpha(t) \neq 0$ para toda t en $[a, b]$ (para que β sea también una curva regular), podemos suponer que $1 - rk_\alpha(t) > 0 \ \forall t \in [a, b]$. Así, la longitud $L_r(\beta)$ se verá como

$$L_r(\beta) = \int_a^b (1 - rk_\alpha(t)) \|\alpha'(t)\| \, dt$$

Si quisiéramos calcular el área del tubo de centro en α y radio R, debemos integrar $L_r(\beta)$ respecto de r desde $-R$ a R. Se tiene

$$\begin{aligned}
\text{Área de } T(\alpha, R) &= \int_{-R}^{R} L_r(\beta) \, dr \\
&= \int_{-R}^{R} \int_a^b (1 - rk_\alpha(t)) \|\alpha'(t)\| \, dt dr \\
&= \int_{-R}^{R} dr \int_a^b \|\alpha'(t)\| \, dt - \int_{-R}^{R} r \, dr \int_a^b k_\alpha(t) \|\alpha'(t)\| \, dt \\
&= 2R \int_a^b \|\alpha'(t)\| \, dt - 0 = 2RL(\alpha)
\end{aligned}$$

en donde $L(\alpha)$ es la longitud de la curva α.

Así, el área del tubo $T(\alpha, R)$ es $2R$ veces la longitud de la curva α.

Ejemplo 1. El camino $\boldsymbol{\alpha}: [0, h] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, 0)$, describe el segmento de recta $0 \leq x \leq h$, $y = 0$. Según el resultado establecido anteriormente, el área del tubo de centro en $\boldsymbol{\alpha}$ y radio R es igual a $2R$ veces la longitud de $\boldsymbol{\alpha}$, que es h. Así pues, según este resultado, el área del tubo es $2Rh$, el cual es efectivamente el área de un rectángulo de dimensiones $h \times 2R$

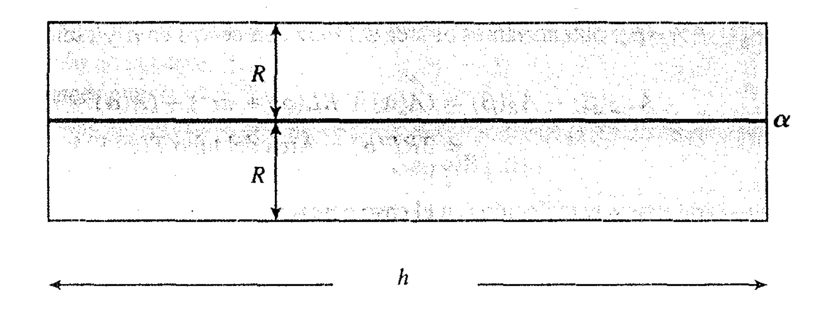


Figura 2. Tubo con centro en $\boldsymbol{\alpha}: [0, h] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, 0)$ y radio R.

Ejemplo 2. Para el camino $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (t, \cosh t)$ el tubo de centro en $\boldsymbol{\alpha}$ y radio R está formado por las curvas r-paralelas a $\boldsymbol{\alpha}$, $-R \leq r \leq R$, $\boldsymbol{\beta}: [a, b] \to \mathbb{R}^2$,

$$
\begin{aligned}
\boldsymbol{\beta}(t) &= \boldsymbol{\alpha}(t) + r\mathbf{n}(t) = (t, \cosh t) + r\frac{1}{\sqrt{1 + \operatorname{senh}^2 t}}(-\operatorname{senh} t, 1) \\
&= \left(t - r\tanh t, \cosh t + \frac{r}{\cosh t}\right)
\end{aligned}
$$

El área de este tubo es $2R$ veces la longitud de $\boldsymbol{\alpha}$ que es $L(\boldsymbol{\alpha}) = \operatorname{senh} b - \operatorname{senh} a$ (ver ejemplo 8 de la sección 5 del capítulo 5). Es decir, el área del tubo es $2R(\operatorname{senh} b - \operatorname{senh} a)$.

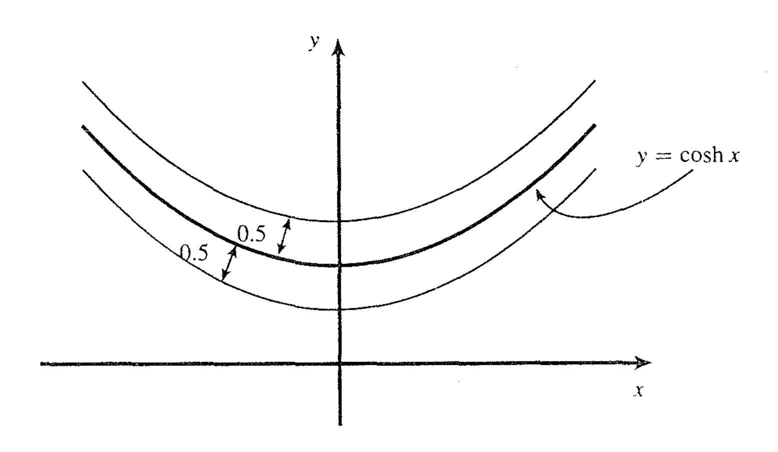


Figura 3. Tubo con centro en la catenaria y radio $r = 0.5$. ■

Ejemplo 3. Si $\boldsymbol{\alpha}$ es una curva cerrada simple positivamente orientada, se vio en el capítulo 5 que el área $A_r(\boldsymbol{\beta})$ que encierra su curva r-paralela $\boldsymbol{\beta}$ (con r tal que $\boldsymbol{\beta}$ siga siendo cerrada simple, por ejemplo, r pequeño) está relacionada con el área $A(\boldsymbol{\alpha})$ que encierra $\boldsymbol{\alpha}$ por

$$A_r(\boldsymbol{\beta}) = A(\boldsymbol{\alpha}) - rL(\boldsymbol{\alpha}) + \pi r^2$$

donde $L(\boldsymbol{\alpha})$ es la longitud de $\boldsymbol{\alpha}$. Si tomamos esta fórmula con $r = \pm R$ y hacemos la resta $A_{-R}(\boldsymbol{\beta}) - A_R(\boldsymbol{\beta})$, obtendríamos el área del tubo con centro en $\boldsymbol{\alpha}$ y radio R. Se tiene

$$\begin{aligned} A_{-R}(\boldsymbol{\beta}) - A_R(\boldsymbol{\beta}) &= (A(\boldsymbol{\alpha}) + RL(\boldsymbol{\alpha}) + \pi r^2) - (A(\boldsymbol{\alpha}) - RL(\boldsymbol{\alpha}) + \pi r^2) \\ &= 2RL(\boldsymbol{\alpha}) = \text{Área del tubo } T(\boldsymbol{\alpha}, R) \end{aligned}$$

como ya habiamos establecido (en el caso general). ■

8.7.2 Tubos en $\mathbb{R}^3$

Para la discusión que se presenta aquí sobre tubos en $\mathbb{R}^3$, necesitamos que la curva $\boldsymbol{\alpha}$ de la que partamos esté parametrizada por longitud de arco y que su curvatura nunca se anule. Sea pues $\boldsymbol{\alpha}: [0, \ell] \to \mathbb{R}^3$ una curva tal (en la que ℓ es la longitud de la curva). Consideremos la superficie $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: S = [0, \ell] \times [0, 2\pi] \subset \mathbb{R}^2 \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = \boldsymbol{\alpha}(u) + r\big((\cos v)\mathbf{N}(u) + (\operatorname{sen} v)\mathbf{B}(u)\big)$$

donde $r > 0$, $\mathbf{N}(u)$ es el vector normal principal y $\mathbf{B}(u)$ es el vector binormal a la curva $\boldsymbol{\alpha}$. En forma geométrica se tiene

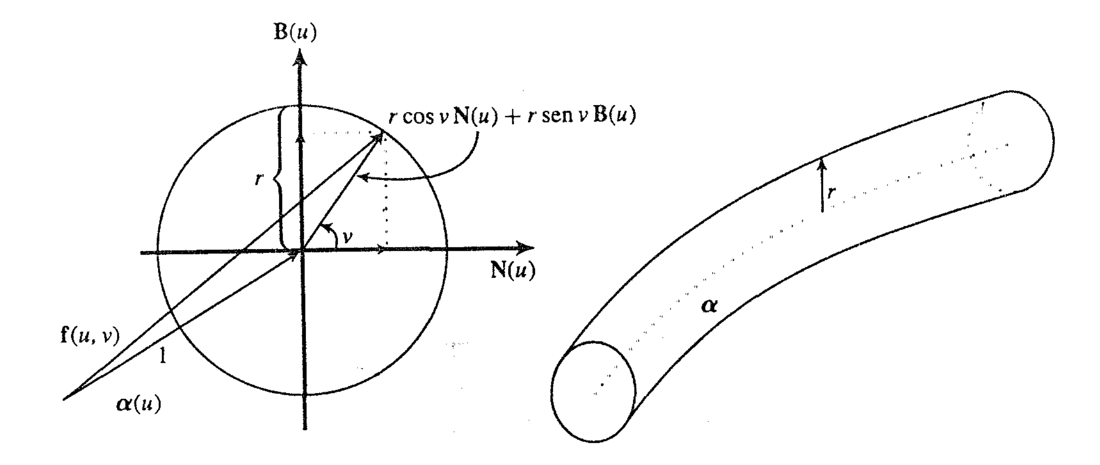


Figura 4. Parametrización de la superficie $K = \mathbf{f}(S)$.

Se ve entonces que los puntos $\mathbf{f}(u, v) \in \mathbb{R}^3$ están en una "superficie tubular" de radio r por cuyo centro va corriendo la curva $\boldsymbol{\alpha}$.

Vamos a calcular el área de esta superficie. Para ello, obtendremos primeramente una expresión para el vector $\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}$. Para realizar estos cálculos ocuparemos en algún momento las fórmulas de Frenet, deducidas en la (parte final de la) sección10 del capítulo 5. Estas son

$$\mathbf{T}'(u) = k(u)\mathbf{N}(u), \quad \mathbf{N}'(u) = -k(u)\mathbf{T}(u) - \tau(u)\mathbf{B}(u), \quad \mathbf{B}'(u) = \tau(u)\mathbf{N}(u)$$

en donde $\mathbf{T}(u) = \boldsymbol{\alpha}'(u)$ es el vector tangente unitario a la curva $\boldsymbol{\alpha}$, $\mathbf{T}'(u)$, $\mathbf{N}'(u)$ y $\mathbf{B}'(u)$ son las derivadas de los vectores tangente $\mathbf{T}(u)$, normal principal $\mathbf{N}(u)$ y binormal $\mathbf{B}(u)$, $k(u)$ es la curvatura y $\tau(u)$ es la torsión de la curva.

Se tiene entonces que

$$\begin{aligned}
\frac{\partial \mathbf{f}}{\partial u} &= \boldsymbol{\alpha}'(u) + r\big((\cos v)\mathbf{N}'(u) + (\operatorname{sen} v)\mathbf{B}'(u)\big) \\
&= \mathbf{T}(u) + r\big((\cos v)(-k(u)\mathbf{T}(u) - \tau(u)\mathbf{B}(u)) + (\operatorname{sen} v)\tau(u)\mathbf{N}(u)\big) \\
&= (1 - r\cos v k(u))\mathbf{T}(u) - r\cos v\tau(u)\mathbf{B}(u) + r\operatorname{sen} v\tau(u)\mathbf{N}(u) \\
\frac{\partial \mathbf{f}}{\partial v} &= r\big(-\operatorname{sen} v\mathbf{N}(u) + \cos v\mathbf{B}(u)\big)
\end{aligned}$$

Entonces

$$\begin{aligned}
\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} &= -r\operatorname{sen} v(1 - rk(u)\cos v)\mathbf{T}(u) \times \mathbf{N}(u) \\
&+ r\cos v(1 - rk(u)\cos v)\mathbf{T}(u) \times \mathbf{B}(u) + r^2\operatorname{sen} v\cos v\ \tau(u)\mathbf{N}(u) \times \mathbf{B}(u) \\
&+ r^2\operatorname{sen} v\cos v\ \tau(u)\mathbf{B}(u) \times \mathbf{N}(u)
\end{aligned}$$

Puesto que $\mathbf{N}(u) \times \mathbf{B}(u) = -\mathbf{B}(u) \times \mathbf{N}(u)$, los dos últimos términos de la expresión anterior se cancelan. Así

$$\mathbf{N_f}(u, v) = -r\operatorname{sen} v(1 - rk(u)\cos v)\mathbf{T}(u) \times \mathbf{N}(u) + r\cos v(1 - rk(u)\cos v)\mathbf{T}(u) \times \mathbf{B}(u)$$

Ciertamente los vectores $\mathbf{T}(u) \times \mathbf{N}(u)$ y $\mathbf{T}(u) \times \mathbf{B}(u)$ son unitarios y ortogonales. Para calcular la norma $\mathbf{N_f}(u, v)$ podemos tomar entonces el teorema de Pitágoras. Se tiene

$$\begin{aligned}
\|\mathbf{N_f}(u, v)\|^2 &= (-r\operatorname{sen} v(1 - rk(u)\cos v))^2\|\mathbf{T}(u) \times \mathbf{N}(u)\|^2 \\
&+ (r\cos v(1 - rk(u)\cos v))^2\|\mathbf{T}(u) \times \mathbf{B}(u)\|^2 \\
&= (1 - rk(u)\cos v)^2 r^2
\end{aligned}$$

de donde

$$\|\mathbf{N_f}(u, v)\| = r(1 - rk(u)\cos v)$$

de modo que el área de la superficie $K = \mathbf{f}(S)$ es

$$\begin{aligned}
\text{Área de } K &= \iint_S \|\mathbf{N_f}(u, v)\|\, du dv = \int_0^\ell \int_0^{2\pi} r(1 - rk(u)\cos v)\, du dv \\
&= r\int_0^\ell du \int_0^{2\pi} dv + r^2 \int_0^\ell k(u)\, du \int_0^{2\pi} \cos v\, dv = 2\pi r\ell
\end{aligned}$$

Así, el área de la superficie tubular de radio r por cuyo centro va la curva $\boldsymbol{\alpha}$, es igual a $2\pi r$ veces la longitud de la curva $\boldsymbol{\alpha}$.

Si consideramos ahora todas las superficies tubulares que llevan por centro a $\boldsymbol{\alpha}$, de radio r, $0 \leq r \leq R$, parametrizadas por la función $\mathbf{f}_r(u, v)$ definidas en $S = [0, \ell] \times [0, 2\pi]$, tenemos que un *tubo* de centro en $\boldsymbol{\alpha}$ y radio R, es

$$T_{\mathbb{R}^3}(\boldsymbol{\alpha}, R) = \{\mathbf{f}_r(S), 0 \leq r \leq R\}$$

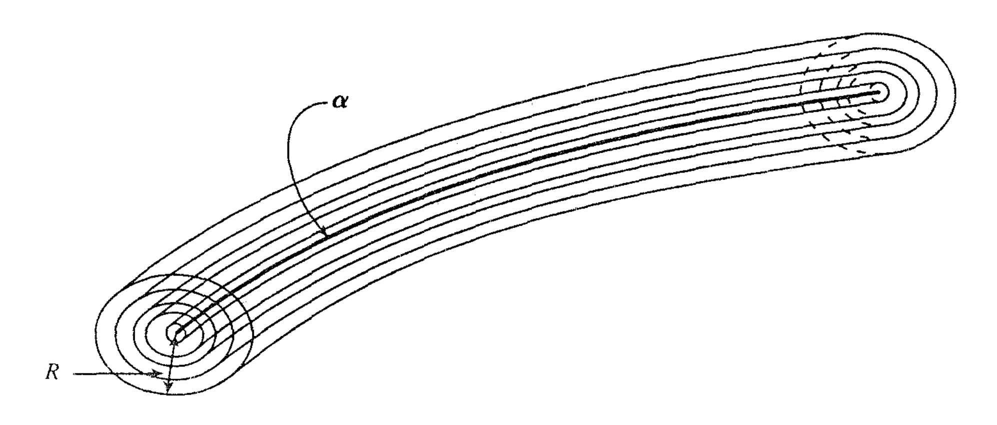


Figura 5. Tubo con centro en $\boldsymbol{\alpha}$ y radio R.

Hemos calculado ya el área de cada *cáscara* $\mathbf{f}_r(S)$ de este tubo. Si quisiéramos calcular su volumen, todo lo que tenemos que hacer es integrar el área de $K = \mathbf{f}_r(S)$ respecto de r, entre 0 y R. Así pues:

$$\text{Volumen de } T_{\mathbb{R}^3}(\boldsymbol{\alpha}, r) = \int_0^R 2\pi r\ell \, dr = \pi R^2 \ell$$

Entonces, el volumen de un tubo en $\mathbb{R}^3$ con centro en $\boldsymbol{\alpha}$ y radio R es πR^2 veces la longitud de la curva $\boldsymbol{\alpha}$.

Ejemplo 4. De nuevo el caso trivial del ejemplo 1: el camino $\boldsymbol{\alpha}: [0, h] \to \mathbb{R}^3$, $\boldsymbol{\alpha}(t) = (t, 0, 0)$ describe el segmento de recta $0 \leq x \leq h$, $y = 0$, $z = 0$. La superficie tubular con centro en $\boldsymbol{\alpha}$ y radio r es un cilindro de longitud h y radio de la base r cuya área (lateral) es, según el resultado que hemos obtenido, $2\pi r$ veces la longitud de $\boldsymbol{\alpha}$, que es h. Es decir, el área lateral del cilindro es $2\pi rh$, como ya sabíamos. También, el volumen del tubo con centro en $\boldsymbol{\alpha}$ y radio r, es el del cilindro mencionado, el cual, según el resultado obtenido, es πr^2 veces la longitud de $\boldsymbol{\alpha}$, que es h. Es decir, el volumen del cilindro es $\pi r^2 h$, como ya también sabíamos. ∎

Si la curva $\boldsymbol{\alpha}: [a, b] \to \mathbb{R}^3$ no está parametrizada por longitud de arco, podemos expresar la parametrización $\mathbf{f}$ de la superficie tubular K, escribiendo las correspondientes expresiones de los

vectores **N** y **B** en términos de $\boldsymbol{\alpha}$ y sus derivadas (ver sección 9 del capítulo 5). Se tiene

$$\mathbf{B} = \frac{\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)}{\|\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)\|}$$

$$\mathbf{N} = \frac{(\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)) \times \boldsymbol{\alpha}'(t)}{\|(\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)) \times \boldsymbol{\alpha}'(t)\|}$$

Entonces la función $\mathbf{f}\colon [a, b] \times [0, 2\pi] \subset \mathbb{R}^2 \to \mathbb{R}^3$ que parametriza la superficie tubular de radio r por cuyo centro va la curva $\boldsymbol{\alpha}$ es

$$\mathbf{f}(u, v) = \boldsymbol{\alpha}(u) + r\left(\cos v \frac{(\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)) \times \boldsymbol{\alpha}'(t)}{\|(\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)) \times \boldsymbol{\alpha}'(t)\|} + \operatorname{sen} v \frac{\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)}{\|\boldsymbol{\alpha}'(t) \times \boldsymbol{\alpha}''(t)\|}\right)$$

Ejemplo 5. Consideremos la hélice $\boldsymbol{\alpha}\colon [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\alpha}(u) = (\cos u, \operatorname{sen} u, 4u)$. Se tiene

$$\begin{aligned}
\boldsymbol{\alpha}'(u) &= (-\operatorname{sen} u, \cos u, 4)\\
\boldsymbol{\alpha}''(u) &= (-\cos u, -\operatorname{sen} u, 0)\\
\boldsymbol{\alpha}'(u) \times \boldsymbol{\alpha}''(u) &= (4 \operatorname{sen} u, -4\cos u, 1)\\
(\boldsymbol{\alpha}'(u) \times \boldsymbol{\alpha}''(u)) \times \boldsymbol{\alpha}'(u) &= (-17\cos u, -17 \operatorname{sen} u, 0)\\
\|\boldsymbol{\alpha}'(u) \times \boldsymbol{\alpha}''(u)\| &= \sqrt{17}\\
\|(\boldsymbol{\alpha}'(u) \times \boldsymbol{\alpha}''(u)) \times \boldsymbol{\alpha}'(u)\| &= 17
\end{aligned}$$

Entonces la superficie tubular de radio r alrededor de la hélice $\boldsymbol{\alpha}$ está parametrizada por $\mathbf{f}\colon [0, 2\pi] \times [0, 2\pi] \subset \mathbb{R}^2 \to \mathbb{R}^3$,

$$\begin{aligned}
\mathbf{f}(u, v) &= (\cos u, \operatorname{sen} u, 4u)\\
&\quad + r\left(\cos v \frac{1}{17}(-17\cos u, -17 \operatorname{sen} u, 0) + \operatorname{sen} v \frac{1}{\sqrt{17}}(4 \operatorname{sen} u, -4\cos u, 1)\right)\\
&= \left(\cos u - r\cos v \cos u + \frac{4}{\sqrt{17}} r \operatorname{sen} v \operatorname{sen} u, \operatorname{sen} u - r\cos v \operatorname{sen} u - \frac{4}{\sqrt{17}} r \operatorname{sen} v \cos u,\right.\\
&\quad \left. 4u + \frac{1}{\sqrt{17}} r \operatorname{sen} v\right)
\end{aligned}$$

Por otra parte, la longitud $L(\boldsymbol{\alpha})$ de la hélice $\boldsymbol{\alpha}$ es

$$L(\boldsymbol{\alpha}) = \int_0^{2\pi} \|\boldsymbol{\alpha}'(t)\|\, dt = \int_0^{2\pi} \sqrt{18}\, dt = 6\sqrt{2}\pi$$

Entonces el área de la superficie tubular de radio r alrededor de la hélice $\boldsymbol{\alpha}$ es $(2\pi r)L(\boldsymbol{\alpha}) = 12\sqrt{2}\pi^2 r$, y el volumen del tubo con centro en $\boldsymbol{\alpha}$ y radio r es $\pi r^2 L(\boldsymbol{\alpha}) = 6\sqrt{2}\pi^2 r^2$.

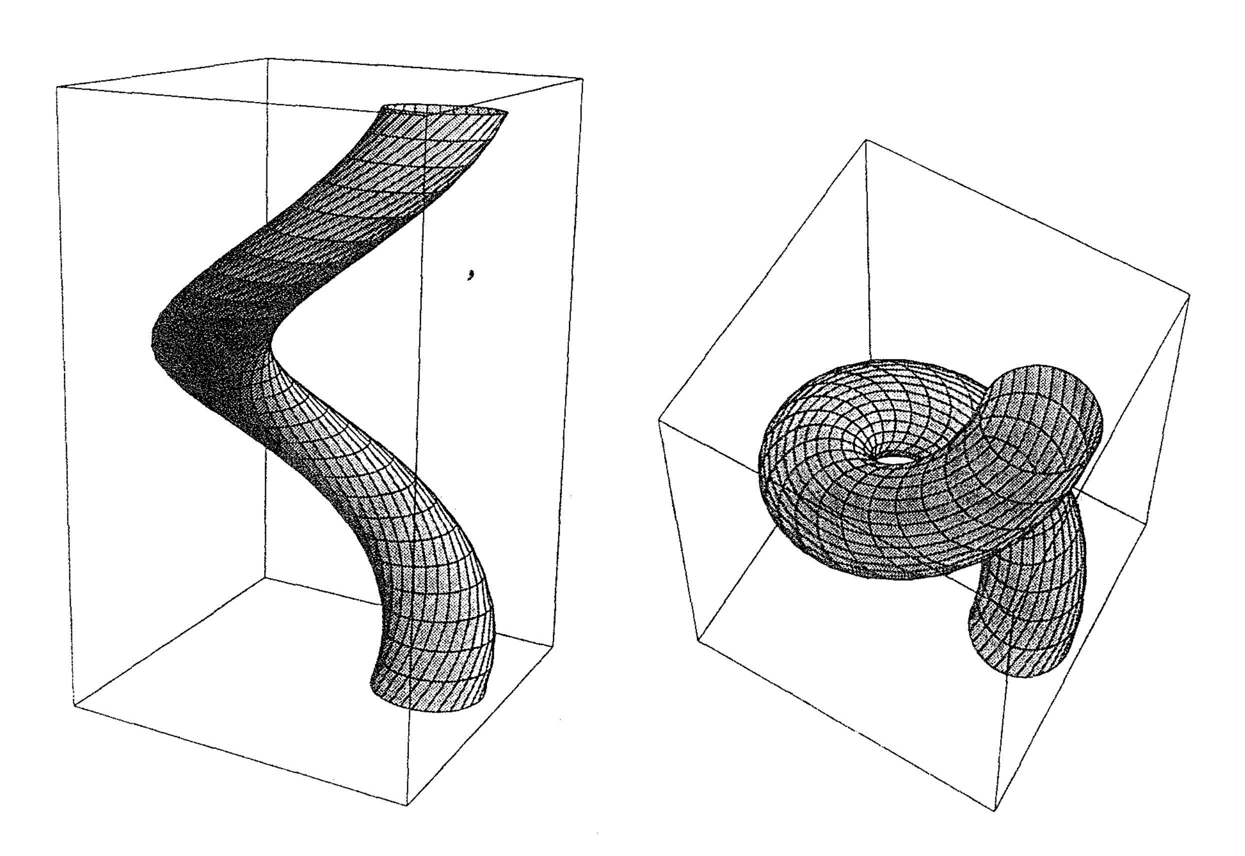


Figura 6. Dos vistas del tubo del ejemplo 5. ∎

Ejercicios (Capítulo 8, Sección 7)

1. Considere la curva $\boldsymbol{\alpha}$: $[0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\alpha}(t) = (r \cos t, r \operatorname{sen} t)$. Calcule el área del tubo con centro en $\boldsymbol{\alpha}$ y radio R.

2. Calcule el área de la superficie de un toro de radios r y R, usando las fórmulas establecidas en esta sección para el área de una superficie tubular. Calcule también el volumen del toro.

Capítulo 9

Integrales de superficie

En los capítulo 6 y 7 hemos generalizado el estudio de integrales que emprendimos en el primer curso de cálculo (donde los integrandos eran funciones reales de una variable real y las regiones de integración eran intervalos de la recta real, o uniones de ellos) considerando en los integrandos de éstas funciones reales de varias variables y/o campos vectoriales, y en las regiones de integración, regiones en el espacio $\mathbb{R}^n$ limitadas por gráficas de funciones de k variables ($k \leq n - 1$ —en las integrales n-múltiples del capítulo 6—, o curvas en el espacio —en las integrales de línea del capítulo 7). En este capítulo emprendemos una nueva generalización del concepto de integral: tendremos en los integrandos funciones reales de varias variables y/o campos vectoriales, y en las regiones de integración, tendremos *superficies* en $\mathbb{R}^3$ (los objetos matemáticos estudiados en el capítulo anterior).

Las integrales que estudiaremos aquí, llamadas *integrales de superficie*, nos servirán para estudiar dos más de los resultados clásicos del cálculo en $\mathbb{R}^n$: el teorema de la divergencia (que en su versión para campos en $\mathbb{R}^2$ ya fue estudiado en el capítulo 7, sección 11, teorema 7.11.1) y el teorema de Stokes, los cuales, junto con el teorema de Green, constituyen los teoremas más importantes del cálculo en $\mathbb{R}^n$ en los casos $n = 2$ y $n = 3$.

En este capítulo se presentará un manejo intenso de los campos vectoriales en $\mathbb{R}^3$, pues ellos aparecerán como integrandos en las integrales de superficie que se ven involucradas en las fórmulas de los teoremas de la divergencia y de Stokes. Estudiaremos los conceptos de divergencia y rotacional de un campo en $\mathbb{R}^3$, los cuales, junto con el concepto de gradiente de una función real definida en algún conjunto U de $\mathbb{R}^3$ (que es un campo en $\mathbb{R}^3$), forman el lenguaje operacional de los campos vectoriales en $\mathbb{R}^3$. En la última sección de este capítulo presentamos una visión conjunta que involucra los conceptos de gradiente, divergencia y rotacional, y que forma parte del estudio clásico del llamado "análisis vectorial".

9.1 Integrales de superficie de funciones reales

En la sección 7 del capítulo 7 se estudiaron las integrales de línea respecto de la longitud de arco: se tenía una función continua $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ y una curva de clase $\mathscr{C}^1$, $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ de modo que $\boldsymbol{\lambda}([a, b]) \subset U$. Entonces la integral de línea de f respecto de la longitud de arco a lo largo de $\boldsymbol{\lambda}$ es

$$\int_{\boldsymbol{\lambda}} f\, ds = \int_a^b f(\boldsymbol{\lambda}(t)) \|\boldsymbol{\lambda}'(t)\|\, dt$$

Los ejemplos que motivaron esta definición (ejemplos 1 y 2) fueron el del pez que come algas en un estanque y el del cálculo de la masa total de un alambre.

Lo que haremos en esta sección será un análogo de la situación anterior, generalizado al caso de superficies en $\mathbb{R}^3$. Para motivar la definición de integral de superficie de una función real, consideremos el problema del cálculo de la masa total de una lámina, cuya forma es la de una superficie simple K. Supongamos que la lámina es muy delgada y que su densidad no sea constante. Podemos entonces pensar en una "función densidad" $\rho: K \to \mathbb{R}$, definida sobre la superficie, de modo que a cada $(x, y, z) \in K$ le asocia el número real $\rho(x, y, z)$, que da el valor (en unidades de masa por unidad de área, digamos gr/cm^2) de la densidad de la lámina en el punto (x, y, z).

Dividamos la superficie K en pequeños rectángulos R_{ij}^K. Si $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ es una función que parametriza a K, cada R_{ij}^K se puede ver cómo la imagen de un rectángulo R_{ij}^S, correspondiente a una partición de la región S en pequeños rectángulos. Si los rectángulos R_{ij}^K en los que quedó dividida la superficie K son pequeños, podemos tener una aproximación de la masa de la lámina en el rectángulo R_{ij}^K como el valor de la densidad ρ en algún punto $(\bar{x}_{ij}, \bar{y}_{ij}, \bar{z}_{ij}) \in R_{ij}^K$, multiplicada por el área de R_{ij}^K. Es decir

$$\text{Masa de la lámina en } R_{ij}^K \approx \rho(\bar{x}_{ij}, \bar{y}_{ij}, \bar{z}_{ij}) \text{ Área de } R_{ij}^K.$$

de modo que

$$\text{Masa total de la lámina } \approx \sum_{i,j} \rho(\bar{x}_{ij}, \bar{y}_{ij}, \bar{z}_{ij}) \text{ Área de } R_{ij}^K$$

Como $R_{ij}^K = \mathbf{f}(R_{ij}^S)$, las áreas de estos rectángulos están relacionadas, según se vio en la sección 6 del capítulo anterior, como

$$\text{Área de } R_{ij}^K = \left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| \text{ Área de } R_{ij}^S$$

donde las derivadas parciales están evaluadas en un punto $(u_{ij}, v_{ij}) \in R_{ij}^S \subset S$. Viendo al punto $(\bar{x}_{ij}, \bar{y}_{ij}, \bar{z}_{ij}) \in R_{ij}^K \subset K$ como la imagen bajo $\mathbf{f}$ del punto $(u_{ij}, v_{ij}) \in R_{ij}^S$, podemos escribir que

$$\text{Masa total de la lámina } \approx \sum_{i,j} \rho(\mathbf{f}(u_{ij}, v_{ij})) \left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| \text{ Área de } R_{ij}^S$$

Esta función corresponde a una suma de Riemann para la función

$$h(u, v) = \rho(\mathbf{f}(u, v)) \left\| \frac{\partial \mathbf{f}}{\partial u}(u, v) \times \frac{\partial \mathbf{f}}{\partial v}(u, v) \right\|$$

sobre la región S. Tomando límite cuando el número de rectángulos de S (y por lo tanto de K) crece al infinito, en tanto que sus áreas tienden a cero, la sumatoria anterior debe tender a la integral doble de la función $h(u, v)$ sobre la región S. Es decir

$$\begin{aligned} \text{Masa total de la lámina } &= \lim \sum_{i,j} \rho(f(u_{ij}, v_{ij})) \left\| \frac{\partial \mathbf{f}}{\partial u}(u_{ij}, v_{ij}) \times \frac{\partial \mathbf{f}}{\partial v}(u_{ij}, v_{ij}) \right\| \text{ Área de } R_{ij}^S \\ &= \iint_S (\rho \circ f)(u, v) \left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| du\,dv \end{aligned}$$

A una expresión como ésta le llamaremos "integral de superficie de la función ρ sobre la superficie simple $K = \mathbf{f}(S)$". En general se tiene la definición siguiente

Definición. Sea $K = \mathbf{f}(S)$ una superficie simple parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$. Sea $\rho: K \to \mathbb{R}$ una función continua definida sobre la superficie K. La integral de superficie de la función ρ sobre K, denotada por $\iint_K \rho\, dA$, se define como

$$\iint_K \rho\, dA = \iint_S \rho(\mathbf{f}(u, v)) \left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| du\, dv$$

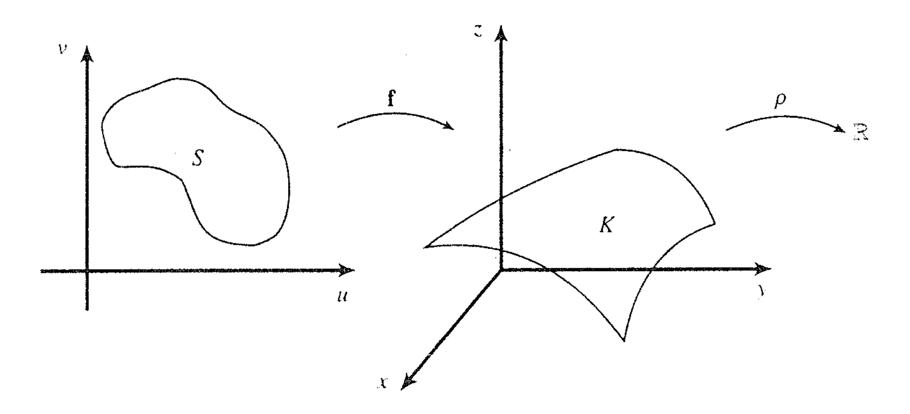


Figura 1. Las funciones $\mathbf{f}$ y ρ, la región S y la superficie K ■

Llamamos la atención al hecho de que si la función $\rho: K \to \mathbb{R}$ que integramos sobre la superficie K, es la función constante $\rho(x, y, z) = 1$, $(x, y, z) \in K$, la integral $\iint_K \rho\, dA$ no es más que la definición de área de la superficie K. Es decir

$$\iint_K 1\, dA = \iint_S \left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| du\, dv = \text{Área de } K$$

(donde la última igualdad es una igualdad numérica. Ver situaciones análogas en el cálculo de áreas de figuras planas con integrales dobles —sección 5.2 capítulo 6—, y volúmenes de cuerpos en $\mathbb{R}^3$ con integrales triples —sección 8.1 capítulo 6).

Veamos algunos ejemplos.

Ejemplo 1. Sea $\rho: \mathbb{R}^3 \to \mathbb{R}$, $\rho(x, y, z) = ax + by + z$, donde a y b son constantes dadas. Consideremos la superficie simple $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}: \{(u, v) | u^2 + v^2 \leq 1\} \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$. K es entonces el pedazo del paraboloide $z = x^2 + y^2$ que se encuentra debajo del plano $z = 1$. Calculemos la integral de superficie de ρ sobre K. Se tiene que

$$\left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| = \|(-2u, -2v, 1)\| = \sqrt{1 + 4(u^2 + v^2)}$$

de modo que

$$\begin{aligned}
\iint_K \rho\, dA &= \iint_S \rho(\mathbf{f}(u,v))\left\|\frac{\partial \mathbf{f}}{\partial u}\times\frac{\partial \mathbf{f}}{\partial v}\right\| du\, dv \\
&= \iint_S (au+bv+u^2+v^2)\sqrt{1+4(u^2+v^2)}du\, dv \overset{\text{polares}}{=\!\!\longrightarrow} \\
&= \int_0^1\int_0^{2\pi}(ar\cos\theta + br\operatorname{sen}\theta + r^2)\sqrt{1+4r^2}\, r\, dr\, d\theta \\
&= 2\pi\int_0^1 \sqrt{1+4r^2}r^3 dr = \frac{\pi}{12}\left(5^{3/2}+\frac{1}{5}\right)
\end{aligned}$$

■

Ejemplo 2. Pensemos en una lámina triangular que tiene la forma de la porción del plano $x+y+z=1$ que se encuentra en el primer octante. Supongamos que la densidad de la lámina en el punto (x, y, z) de ella es proporcional a la distancia del punto al origen, siendo igual a 1 en el punto $(0, 0, 1)$. Se quiere determinar la masa total de la lámina. Una función $\mathbf{f}$ que parametriza a la superficie K (el plano $x+y+z=1$ en el primer octante) es $\mathbf{f}: \{(u,v)|0\le u\le 1, 0\le v\le 1-u\}\subset\mathbb{R}^2\to\mathbb{R}^3$, $\mathbf{f}(u,v)=(u,v,1-u-v)$. Para $(x,y,z)=\mathbf{f}(u,v)\in K$, la función densidad es $\rho(x,y,z)=c(x^2+y^2+z^2)$ donde $c=1$ (pues $\rho(0,0,1)=1=c(0+0+1)$). Se trata entonces de calcular la integral de la función ρ sobre la superficie K. Se tiene

$$\begin{aligned}
\iint_K \rho\, dA &= \iint_S \rho(\mathbf{f}(u,v))\left\|\frac{\partial \mathbf{f}}{\partial u}\times\frac{\partial \mathbf{f}}{\partial v}\right\| du\, dv \\
&= \int_S\int (u^2+v^2+(1-u-v)^2)\|(1,1,1)\|du\, dv \\
&= \sqrt{3}\int_0^1\int_0^{1-u}(u^2+v^2+(1-u-v)^2)du\, dv = \frac{\sqrt{3}}{4}
\end{aligned}$$

■

El valor de la integral de superficie $\iint_K \rho\, dA$ no depende de la parametrización concreta que se tenga de K. En efecto, si $\mathbf{g}=\mathbf{f}\circ\boldsymbol{\varphi}: S'\subset\mathbb{R}^2\to\mathbb{R}^3$ es una reparametrización de K, donde $\boldsymbol{\varphi}: S'\to S$, $\boldsymbol{\varphi}=(\varphi_1,\varphi_2)$ es entonces una biyección de clase $\mathscr{C}^1$ con determinante jacobiano $\frac{\partial(\varphi_1,\varphi_2)}{\partial(s,t)}$ nunca nulo, se tiene

$$\begin{aligned}
\iint_K \rho\, dA &\underset{\text{según } f}{=} \iint_S \rho(\mathbf{f}(u,v))\left\|\frac{\partial \mathbf{f}}{\partial u}\times\frac{\partial \mathbf{f}}{\partial v}\right\| du\, dv \xrightarrow[(u,v)=\varphi(s,t)]{\text{cambio de variable}} \\
&= \iint_{S'} \rho(f(\varphi(s,t)))\|N_{\mathbf{f}}(\varphi(s,t))\|\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(s,t)}\right| ds\, dt \\
&= \iint_{S'} \rho(g(s,t))\|N_{\mathbf{g}}(s,t)\|ds\, dt \underset{\text{según } g}{=} \int_K \rho\, dA
\end{aligned}$$

en donde se usó nuevamente la fórmula

$$\frac{\partial \mathbf{g}}{\partial s}\times\frac{\partial \mathbf{g}}{\partial t} = \mathbf{N}_{\mathbf{g}}(s,t) = \frac{\partial(\varphi_1,\varphi_2)}{\partial(s,t)}\mathbf{N}_{\mathbf{f}}(u,v) = \frac{\partial(\varphi_1,\varphi_2)}{\partial(s,t)}\frac{\partial \mathbf{f}}{\partial u}\times\frac{\partial \mathbf{f}}{\partial v}$$

establecida en la sección 2 del capítulo 8.

Ejemplo 3. Sea $\mathbf{g}: S = \{(s,t)|s^2 + t^2 \leq 1/2\} \subset \mathbb{R}^2 \to \mathbb{R}^3$, la función $\mathbf{g}(s,t) = (s+t, s-t, 2s^2 + 2t^2)$. Esta es una reparametrización de la superficie parabólica K del ejemplo 1 (ver también ejemplo 5, sección 6, capítulo 8). Calculemos la integral de superficie de la misma función $\rho(x,y,z) = ax + by + z$ del ejemplo 1. Se tiene

$$\begin{aligned}\iint_K \rho\, dA &= \iint_S \rho(\mathbf{g}(s,t))\left\|\frac{\partial \mathbf{g}}{\partial s} \times \frac{\partial \mathbf{g}}{\partial t}\right\| ds\, dt \\ &= \iint_S (a(s+t) + b(s-t) + 2s^2 + 2t^2)\|(4s+4t, 4s-4t, -2)\| ds\, dt \\ &= \iint_S ((a+b)s + (a-b)t + 2s^2 + 2t^2)\sqrt{4 + 32(s^2+t^2)}\, ds\, dt \overset{\text{polares}}{=\!\!\longrightarrow} \\ &= \int_0^{2\pi}\int_0^{1/\sqrt{2}} ((a+b)r\cos\theta + (a-b)r \operatorname{sen}\theta + 2r^2)\sqrt{4+32r^2}\, r\, dr\, d\theta \\ &= \frac{\pi}{12}\left(5^{3/2} + \frac{1}{5}\right)\end{aligned}$$

que es el mismo resultado que se obtuvo en el ejemplo 1. ■

Nuevamente advertimos que la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ que parametriza a la superficie K puede perder sus características (como la inyectividad) en un subconjunto $\tilde{S}$ de S de medida nula siguiendo sin cambios la definición de integral de superficie de una función continua $\rho: K \to \mathbb{R}$ sobre K.

Ejemplo 4. Sea $K = \mathbf{f}(S)$ la esfera unitaria, parametrizada por $\mathbf{f}: [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$

$$\mathbf{f}(u,v) = (\cos u \operatorname{sen} v, \operatorname{sen} u \operatorname{sen} v, \cos v)$$

Como sabemos, esta función no es inyectiva en los puntos de la frontera del rectángulo $[0, 2\pi] \times [0, \pi]$ (ver ejemplo 2 de la sección 4 del capítulo 8). Esto no impide, sin embargo, que calculemos la integral de superficie de la función continua $\rho: K \to \mathbb{R}$, $\rho(x,y,z) = \frac{1}{\sqrt{x^2+y^2+(z-a)^2}}$ (donde a es un número dado $0 < a < 1$) sobre K. Se tiene

$$\begin{aligned}\iint_K \rho\, dA &= \iint_S \rho(\mathbf{f}(u,v))\left\|\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right\| du\, dv \\ &= \iint_S \frac{1}{\sqrt{(\cos u \operatorname{sen} v)^2 + (\operatorname{sen} u \operatorname{sen} v)^2 + (\cos v - a)^2}} (\operatorname{sen} v) du\, dv \\ &= \iint_S \frac{\operatorname{sen} v du\, dv}{\sqrt{\operatorname{sen}^2 v + (\cos v - a)^2}} = \int_0^{2\pi} du \int_0^{\pi} \frac{\operatorname{sen} v\, dv}{\sqrt{1 + a^2 - 2a\cos v}} = 4\pi\end{aligned}$$ ■

El concepto de integral de superficie dado para superficies simples, puede ser extendido al caso de superficies más generales. Dada una superficie $K = \mathbf{f}(S_1 \cup S_2 \cup \ldots \cup S_k)$ en que $S_1, S_2, \ldots, S_k$

son regiones del plano del tipo I y II (según se dio la definición de este concepto en la sección 4 del capítulo 8), la integral de superficie de la función continua $\rho: K \to \mathbb{R}$ sobre la superficie K es

$$\iint_K \rho\, dA = \sum_{i=1}^{k} \iint_{S_i} \rho(\mathbf{f}(u,v)) \left\| \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} \right\| du\, dv$$

9.1.1 Aplicaciones (I): Valor medio de una función definida en una superficie

Sea K una superficie y $\rho: K \to \mathbb{R}$ una función continua definida en K. Se define el *valor medio de* ρ *sobre* K, denotado por $\bar{\rho}_K$, como

$$\bar{\rho}_K = \frac{\displaystyle\iint_K \rho\, dA}{\displaystyle\iint_K dA} = \frac{1}{\text{Área de } K} \iint_K \rho\, dA$$

Así, el valor $\bar{\rho}_K$ es una estimación del "promedio de los valores de ρ sobre K". Más aún, el valor $\bar{\rho}_K$ es tomado por la función ρ en un punto $\mathbf{q} \in K$. De esta manera, se puede decir que siendo $\rho: K \to \mathbb{R}^3$ una función continua definida en la superficie K, existe un punto $\mathbf{q} \in K$ tal que

$$\iint_K \rho\, dA = \rho(\mathbf{q}) \iint_K dA = \bar{\rho}_K \iint_K dA = (\text{ Valor medio de } \rho \text{ en } K)(\text{ Área de } K)$$

Este resultado es conocido como "Teorema del Valor Medio" (en su versión para funciones continuas definidas en superficies). Obsérvese que se trata de la misma idea involucrada en los "Teoremas del Valor Medio" establecidos para funciones reales de una variable real (visto en el primer curso de cálculo), para funciones de n variables sobre una región K de $\mathbb{R}^n$ (visto en las secciones 6.5.4 y 6.8.3 del capítulo 6) y para funciones de n variables sobre una curva (visto en la sección 7.7.2 del capítulo 7).

Ejemplo 5. Consideremos de nuevo la lámina del ejemplo 2, situada en la porción del plano $x + y + z = 1$ que se encuentra en el primer octante. En tal ejemplo, se calculó la masa total de la lámina, siendo ésta igual a $\frac{\sqrt{3}}{4}$ (unidades de masa). Este valor es el de la integral de la función densidad $\rho(x, y, z) = x^2 + y^2 + z^2$ sobre la superficie K (donde se encuentra la lámina). Puesto que el área de K es $\frac{\sqrt{3}}{2}$, la *densidad media* de la lámina es

$$\bar{\rho}_K = \frac{\displaystyle\iint_K \rho\, dA}{\displaystyle\iint_K dA} = \frac{\text{Masa total de la lámina}}{\text{área de la lámina}} = \frac{\frac{\sqrt{3}}{4}}{\frac{\sqrt{3}}{2}} = \frac{1}{2}$$ ■

Ejemplo 6. Calculemos el valor medio de la función $\rho: \mathbb{R}^3 - \{(0, 0, a)\} \to \mathbb{R}$ dada por $\rho(x, y, z) = \frac{1}{\sqrt{x^2+y^2+(z-a)^2}}$, en donde $0 < a < 1$, sobre la esfera unitaria $x^2 + y^2 + z^2 = 1$.

Sea $\mathbf{f}: [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (\cos u \operatorname{sen} v, \operatorname{sen} u \operatorname{sen} v, \cos v)$ la función que parametriza a la esfera K. El valor de la integral de ρ sobre K ya se calculó en el ejemplo 4. Este es 4π. Como el área de K es 4π se tiene que el valor medio de ρ sobre K es

$$\bar{\rho}_K = \frac{\iint_K \rho\, dA}{\iint_K dA} = \frac{4\pi}{4\pi} = 1$$

El resultado aquí obtenido tiene una interpretación física interesante: dado un punto $\mathbf{p} = (x, y, z) \in \mathbb{R}^3$ y una carga unitaria $\mathbf{q}_0$ en el punto $\mathbf{q}$ fijo, el potencial de $\mathbf{p}$ debido a la carga $\mathbf{q}_0$ es igual a $\frac{1}{d}$, donde d es la distancia entre $\mathbf{p}$ y $\mathbf{q}$. Obsérvese que la función ρ considerada en este ejemplo, $\rho(x, y, z) = \frac{1}{\sqrt{x^2+y^2+(z-a)^2}}$, da justamente el valor del potencial de $\mathbf{p} = (x, y, z)$ debido a una carga unitaria $\mathbf{q}_0$ que se encuentra en el punto $(0, 0, a)$. Lo que hicimos en este ejemplo fue entonces calcular el *potencial medio* de los puntos de la esfera unitaria debido a una carga unitaria $\mathbf{q}_0$ que se encuentra *dentro de la esfera* (¿por qué?). Nótese que el valor obtenido *no* depende de la posición de la carga $\mathbf{q}_0$. Así, el potencial medio de los puntos de la esfera unitaria debido a una carga unitaria que se encuentra *en cualquier* punto de su interior, es igual a 1. ■

9.1.2 Aplicaciones (II): Centros de masa y momentos de superficies

Sea $K = \mathbf{f}(S)$ una superficie y sea $\rho: K \to \mathbb{R}$ una función que da la densidad de (una lámina situada sobre) la superficie K. Es decir, $\rho(x, y, z) =$ densidad (por ejemplo en gr/cm^2) de la superficie K en el punto $(x, y, z) \in K$. El centro de masa de K se localiza en el punto $(\bar{x}, \bar{y}, \bar{z})$ donde

$$\bar{x} = \frac{\iint_K \rho(x, y, z) x\, dA}{\iint_K \rho(x, y, z)\, dA}, \quad \bar{y} = \frac{\iint_K \rho(x, y, z) y\, dA}{\iint_K \rho(x, y, z)\, dA}, \quad \bar{z} = \frac{\iint_K \rho(x, y, z) z\, dA}{\iint_K \rho(x, y, z)\, dA}$$

Después de haber visto fórmulas análogas para figuras planas (sección 6.5.3, capítulo 6), cuerpos en el espacio (sección 8.2, capítulo 6) y curvas en el plano (sección 7.7.2, capítulo 7), las fórmulas anteriores deben parecer naturales. Obsérvese que su denominador es la masa total de la lámina que se encuentra en K, y sus numeradores son los *momentos estáticos* de la superficie K respecto de los planos yz, xz y xy, respectivamente. Si la densidad de la lámina es constante (si la lámina es homogénea), las fórmulas anteriores toman el aspecto más simple

$$\bar{x} = \frac{\iint_K x\, dA}{\iint_K dA}, \quad \bar{y} = \frac{\iint_K y\, dA}{\iint_K dA}, \quad \bar{z} = \frac{\iint_K z\, dA}{\iint_K dA}$$

en las que ahora el denominador es el área de K y los numeradores son los valores medios de las funciones $\varphi_1(x, y, z) = x$, $\varphi_2(x, y, z) = y$, $\varphi_3(x, y, z) = z$ (respectivamente), sobre K.

Ejemplo 7. Calculemos las coordenadas del centro de masa de la superficie parabólica $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $S = \{(u, v)|u^2+v^2 \leq 1\}$, $\mathbf{f}(u, v) = (u, v, u^2+v^2)$. Supongamos que se trata de un cuerpo homogéneo. El área de $K = \mathbf{f}(S)$ es $\frac{\pi}{6}\left(5^{3/2} - 1\right)$ (calculada en el ejemplo 1 de la sección 6 del capítulo 8). Por razones de simetría (y de la homogeneidad de la superficie), podemos asegurar que $\bar{x} = \bar{y} = 0$. Obtengamos entonces el numerador de la coordenada $\bar{z}$. Se tiene

$$\begin{aligned}
\iint_K z\, dA &= \iint_S (u^2+v^2)\left\|\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right\| du\, dv \\
&= \iint_S (u^2+v^2)\sqrt{1+4u^2+4v^2}\, du\, dv \overset{\text{polares}}{=\!=\!\longrightarrow} \\
&= \int_0^{2\pi}\int_0^1 \sqrt{1+4r^2}\, r^3 dr\, d\theta = \frac{\pi}{12}\left(5^{3/2} + \frac{1}{5}\right)
\end{aligned}$$

Entonces la coordenada $\bar{z}$ es:

$$\bar{z} = \frac{\iint_K z\, dA}{\iint_K dA} = \frac{\frac{\pi}{12}\left(5^{3/2} + \frac{1}{5}\right)}{\frac{\pi}{6}(5^{3/2} - 1)} = \frac{1}{620}(15\sqrt{5} + 313)(\approx 0.5589)$$ ■

Ejemplo 8. Calculemos el centro de masa de la superficie parabólica del ejemplo anterior a la que le hemos añadido "la tapa". Considere la función $\mathbf{f}: S_1 \cup S_2 \to \mathbb{R}^3$, donde

$$S_1 = \{(u, v)|u^2 + v^2 \leq 1\},\ S_2 = \{(u, v)|(u-3)^2 + (v-3)^2 \leq 1\}$$

dada por

$$\mathbf{f}(u, v) = \begin{cases} (u, v, u^2+v^2) & \text{si } (u, v) \in S_1 \\ (u-3, v-3, 1) & \text{si } (u, v) \in S_2 \end{cases}$$

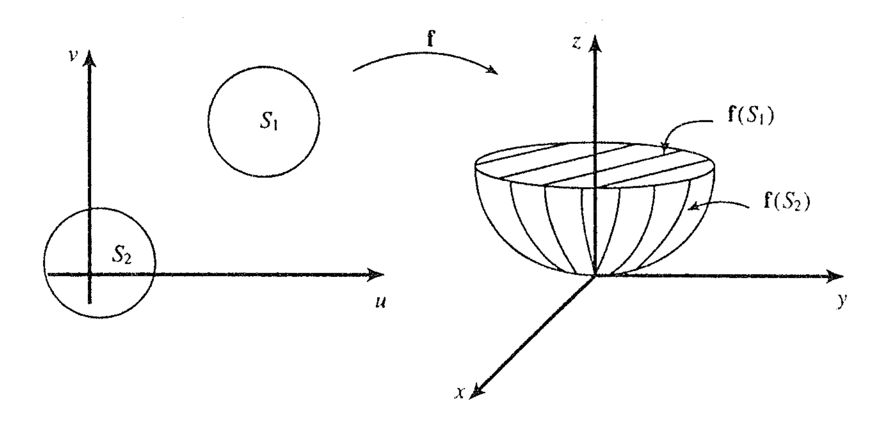


Figura 2. La superficie parabólica del ejemplo 8



Se verifica que $K = \mathbf{f}(S)$ es una superficie compuesta por la porción del paraboloide $z = x^2 + y^2$ que se encuentra por debajo del plano $z = 1$ ésta es $\mathbf{f}(S_1)$) y el plano mismo $z = 1$ (o sea es $\mathbf{f}(S_2)$). El área de K es la suma de las áreas de $\mathbf{f}(S_1)$ y $\mathbf{f}(S_2)$. Es decir, con el cálculo del ejemplo 1 de la sección 6 del capítulo 8, se tiene

$$\text{Área de } K = \frac{\pi}{6}\left(5^{3/2} - 1\right) + \pi$$

Por razones de simetría es claro que $\bar{x} = \bar{y} = 0$. El numerador de $\bar{z}$ es (usando el cálculo del ejemplo anterior)

$$\begin{aligned}\iint_K z\,dA &= \iint_{S_1} (u^2 + v^2)\left\|\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right\| du\,dv + \iint_{S_2} (1)\|(0, 0, 1)\|du\,dv \\ &= \frac{\pi}{12}\left(5^{3/2} + \frac{1}{5}\right) + \text{área de } S_2 = \frac{\pi}{12}\left(5^{3/2} + \frac{1}{5}\right) + \pi\end{aligned}$$

Entonces

$$\bar{z} = \frac{\displaystyle\iint_K z\,dA}{\displaystyle\iint_K dA} = \frac{\frac{\pi}{12}\left(5^{3/2} + \frac{1}{5}\right) + \pi}{\frac{\pi}{6}(5^{3/2} - 1) + \pi} = \frac{1}{50}(9\sqrt{5} + 16)\ (\approx 0.7225)$$

Es decir, el centro de masa de esa superficie se encuentra en el punto $\left(0, 0, \frac{1}{50}(9\sqrt{5} + 16)\right) \approx (0, 0, 0.7225)$. Compare este resultado con el del ejemplo anterior, en el que hallamos que el centro de masa de la porción del paraboloide está en $\left(0, 0, \frac{1}{620}(15\sqrt{5} + 313)\right) \approx (0, 0, 0.5589)$. Compare también con el resultado del ejemplo 8 de la sección 8 del capítulo 6, en el que hallamos que el centro de masa de la región en $\mathbb{R}^3$ que ocupa el paraboloide $z = x^2 + y^2$ por debajo del plano $z = 1$, se encontraba en $\left(0, 0, \frac{2}{3}\right)$. ■

Ejemplo 9. Consideremos una "porción de toro" $K = \mathbf{f}(S)$, parametrizada por $\mathbf{f}\colon [0, u_0] \times [0, 2\pi] \to \mathbb{R}^3$

$$\mathbf{f}(u, v) = ((R + r\cos v)\cos u, (R + r\cos v)\operatorname{sen} u, r\operatorname{sen} v))$$

(ver ejemplo 6 de la sección 4 del capítulo 8), como se muestra en la figura 3

Se tiene

$$\left\|\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right\| = r(R + r\cos v)$$

(ver ejemplo 9 de la sección 6 del capítulo 8). Suponiendo que K es una superficie homogénea, calculemos su centro de masa. Por razones de simetría, es claro que la coordenada $\bar{z}$ será igual a 0. El área de K es igual a

$$\begin{aligned}\text{Área de } K &= \iint_K dA = \iint_S \left\|\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right\| du\,dv \\ &= \int_0^{u_0}\int_0^{2\pi} r(R + r\cos v)du\,dv = 2\pi r R u_0\end{aligned}$$

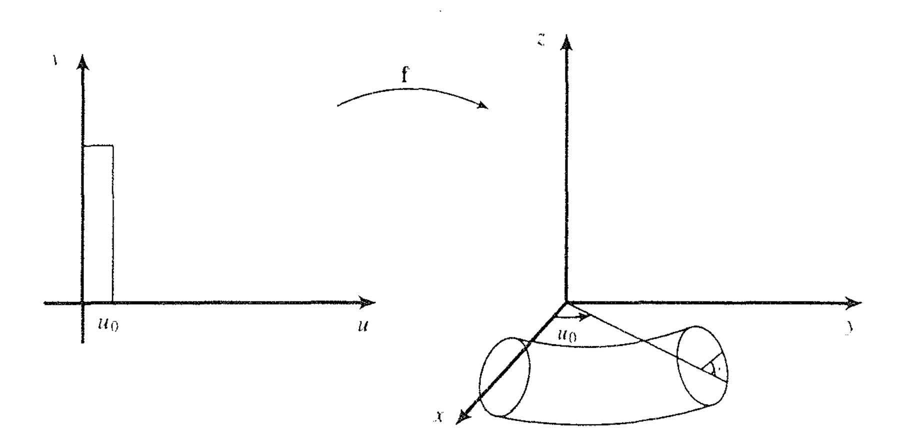


Figura 3. La porción de toro del ejemplo 9.

El numerador de la coordenada $\bar{x}$ es

$$\iint_K x\,dA = \iint_S (R + r\cos v)\cos u(r(R + r\cos v))du\,dv$$

$$= r\int_0^{u_0} \cos u\,du \int_0^{2\pi} (R + r\cos v)^2 dv = \pi r(r^2 + 2R^2)\operatorname{sen} u_0$$

y el numerador de la coordenada $\bar{y}$ es

$$\iint_K y\,dA = \iint_S (R + r\cos v)\operatorname{sen} u(r(R + r\cos v))du\,dv$$

$$= r\int_0^{u_0} \operatorname{sen} u\,du \int_0^{2\pi} (R + r\cos v)^2 dv = \pi r(r^2 + 2R^2)(1 - \cos u_0)$$

Entonces

$$\bar{x} = \frac{\displaystyle\iint_K x\,dA}{\displaystyle\iint_K dA} = \frac{r^2 + 2R^2}{2Ru_0}\operatorname{sen} u_0$$

$$\bar{y} = \frac{\displaystyle\iint_K y\,dA}{\displaystyle\iint_K dA} = \frac{r^2 + 2R^2}{2Ru_0}(1 - \cos u_0)$$

Es decir, el centro de masa de la superfcie K se encuentra en

$$\frac{r^2 + 2R^2}{2Ru_0}\left(\operatorname{sen} u_0, 1 - \cos u_0, 0\right)$$

Si $u_0 = 2\pi$, la superficie K es el toro completo y su centro de masa se encuentra en $(0, 0, 0)$, como tenía que ocurrir. ■

Ejercicios (Capítulo 9, Sección 1)

1. Sea K una superficie simple y sean $\rho_1, \rho_2 \colon K \to \mathbb{R}$ dos funciones continuas definidas en K. Demuestre que

$$\iint_K (\rho_1 + c\rho_2)\, dA = \iint_K \rho_1\, dA + c \iint_K \rho_2\, dA$$

donde c es un número real.

En los ejercicios 2—12, calcule la integral de superficie de la función $\rho \colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ dada sobre la superficie indicada.

2. $\rho(x, y, z) = x$, sobre $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, x \geq 0, z \geq 0\}$.
3. $\rho(x, y, z) = y$, sobre $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, x \geq 0, z \geq 0\}$.
4. $\rho(x, y, z) = z$, sobre $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, x \geq 0, z \geq 0\}$.
5. $\rho(x, y, z) = 3x + 2y - 6z$, sobre $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, x \geq 0\ z \geq 0\}$. (Sugerencia: use los resultados de los 4 ejercicios anteriores).
6. $\rho(x, y, z) = 2x + y$, sobre $K = \{(x, y, z)|x^2 + y^2 = 2, -1 \leq z \leq 1\}$.
7. $\rho(x, y, z) = x - y + z$, sobre $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1, z \leq 0\}$.
8. $\rho(x, y, z) = xy + xz$, sobre $K = \{(x, y, z)|x + y + z = 1, x \geq 0, y \geq 0, z \geq 0\}$.
9. $\rho(x, y, z) = x^2y$, sobre la porción del plano $2x + 3y - 5z = 1$ que se encuentra en el primer octante.
10. $\rho(x, y, z) = 4$, sobre el cubo $K = [a, b] \times [c, d] \times [e, f]$.
11. $\rho(x, y, z) = 5$, sobre la parte de la superficie $|x| + |y| + |z| = 1$ que se encuentra en el primero y segundo octantes.
12. $\rho(x, y, z) = xyz$, sobre la porción del paraboloide $z = 1 - x^2 - y^2$ que se encuentra por encima del plano $z = 0$.
13. Calcule la integral de superficie de la función $\rho(x, y, z) = x^2 + y^2$, sobre el cilindro $K = \{(x, y, z)|x^2 + y^2 = 1, a \leq z \leq b\}$. (Sugerencia: observe que la función ρ *es constante* en K). ¿Cuál es el valor medio de la función ρ sobre la superficie K? Explique
14. Calcule la integral de superficie de la función $\rho(x, y, z) = x^2 + y^2 + z^2$ sobre la esfera $K = \{(x, y, z)|x^2 + y^2 + z^2 = c^2\}$. (Sugerencia: la función ρ es constante en K). ¿Cuál es el valor medio de la función ρ sobre la superficie K? Explique.

15. Sea $K = \{(x, y, z)|x^2+y^2+z^2 = c^2\}$. Considere las funciones $\rho_1(x, y, z) = x^2$, $\rho_2(x, y, z) = y^2$, $\rho_3(x, y, z) = z^2$. Sin hacer cálculos, argumente por qué

$$\iint_K \rho_1\, dA = \iint_K \rho_2\, dA = \iint_K \rho_3\, dA$$

Con el resultado del ejercicio anterior (junto con el del ejercicio 1) calcule el valor de estas integrales.

16. Calcule el valor medio de la función $\rho(x, y, z) = x$, sobre la porción del plano $z = x$ cuya proyección en el plano xy es el cuadrado $[-1, 1] \times [-1, 1]$. Interprete su resultado.

17. Repita el ejercicio anterior con la función $\rho(x, y, z) = y$, y el plano $z = y$.

18. Repita los dos ejercicios anteriores si consideramos la superficie K como la porción del plano $z = x$ (en el ejercicio 16) o del plano $z = y$ (en el ejercicio 17), cuya proyección en el plano xy es $R = \{(x, y)|x^2 + y^2 \leq 1\}$. ¿Qué sucede si consideramos, más en general, la porción de estos planos cuya proyección en el plano xy es una región R simétrica respecto del origen de coordenadas? Explique.

19. ¿Cuál es el valor promedio del producto de tres números no negativos cuando cada uno de estos varía de tal manera que su suma siempre es igual a la unidad?

20. ¿Cuál es el valor promedio de la suma de tres números no negativos cuando cada uno de estos varía de tal modo que la suma de sus cuadrados es siempre igual a la unidad?

21. ¿Cuál es el valor promedio de la suma de los cuadrados de tres números no negativos cuando cada uno de estos varía de tal modo que su suma es siempre igual a la unidad?

22. Calcule la masa total de una superficie en forma de cilindro circular recto de radio de la base R y altura H (sin tapas), si la densidad en cada punto de él es (numéricamente) igual a la distancia del punto a la base del cilindro.

23. Determine las coordenadas del centro de masa del casquete esférico $x^2 + y^2 + z^2 = 1$, $z \geq 0$ (se supone que la superficie es homogénea).

24. Determine las coordenadas del centro de masa de la porción de la esfera unitaria correspondiente a la región (en coordenadas esféricas) $R = \{(r, \theta, \phi)|r = 1, 0 \leq \theta \leq \theta_0, 0 \leq \phi \leq \pi\}$, en donde $0 \leq \theta_0 \leq 2\pi$, (se supone que la superficie es homogénea).

25. Determine las coordenadas del centro de masa de la superficie homogénea en forma de cono circular recto de radio de la base R y altura H: **a.** sin incluír la base; **b.** incluyendo la base.

9.2 Integrales de superficie de campos vectoriales

En esta sección introducimos las integrales de superficie para campos vectoriales en $\mathbb{R}^3$. La idea física que hay detrás de este concepto está relacionada con el "flujo" de un campo a través de una superficie K. En forma más precisa, supongamos que $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ es un campo (continuo) en $\mathbb{R}^3$ y que K es una superficie simple. Si $\mathbf{F}$ representa el campo de velocidades de un fluido (digamos que $\mathbf{F}(\mathbf{q})$ está dado en m/seg), se trata de ver cuál es el flujo (digamos en m^3/seg) de éste a través de la superficie K.

Antes de ver la definición correspondiente en la que hallaremos respuesta a esta pregunta, hagamos algunas operaciones breves que nos permitirán ver la definición mencionada de manera natural. Nótese que para cada $\mathbf{q} \in K$, el vector $\mathbf{F}(\mathbf{q})$ puede ser descompuesto como una suma de vectores $F^{(1)}(\mathbf{q})$ y $F^{(2)}(\mathbf{q})$, en los que $F^{(1)}(\mathbf{q})$ se encuentra en el plano tangente a K en $\mathbf{q}$ y $F^{(2)}(\mathbf{q})$ es ortogonal a $F^{(1)}(\mathbf{q})$, es decir, se encuentra en "la normal" a K en $\mathbf{q}$. De estas dos componentes, es claro que la que proporciona información sobre el flujo de $\mathbf{F}$ a través de K es $F^{(2)}(\mathbf{q})$, pues ésta es la que, por decirlo así, está atravesando la superficie.

Es importante mencionar que las superficies que se considerarán en esta sección son superficies orientables en las que hay un campo continuo de vectores normales $\mathbf{N}: K \to \mathbb{R}^3$ que le da una orientación determinada a la superficie. Obsérvese entonces que, para cada $\mathbf{q} \in K$, los vectores $F^{(2)}(\mathbf{q})$ y $\mathbf{N}(\mathbf{q})$ se encuentran en una misma línea (la normal a K en $\mathbf{q}$). La idea intuitiva que manejaremos es que, siendo $\mathbf{N}(\mathbf{q})$ un vector que apunta hacia uno de los lados de K, digamos que a su "exterior", cuando $F^{(2)}(\mathbf{q})$ y $\mathbf{N}(\mathbf{q})$ coinciden en dirección y sentido, diremos que el flujo está "saliendo de K": en el caso contrario, diremos que el flujo está "entrando a K".

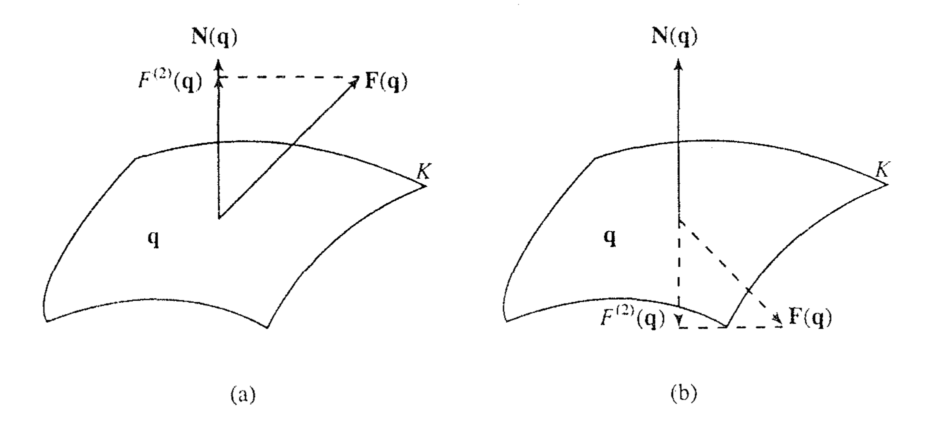


Figura 1. (a) El flujo sale de K. (b) El flujo entra a K.

Puesto que $\mathbf{N}(\mathbf{q})$ es un vector unitario, la componente $F^{(2)}(\mathbf{q})$ se encuentra como

$$\text{componente de } \mathbf{F}(\mathbf{q}) \text{ en } \mathbf{N}(\mathbf{q}) = \mathbf{F}(\mathbf{q}) \cdot \mathbf{N}(\mathbf{q})$$

(ver sección 2 del capítulo 1).

Sea entonces $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ una parametrización de la superficie K, y tomemos la orientación de K según el campo continuo de vectores normales $\mathbf{N}: K \to \mathbb{R}^3$. Si hacemos una partición en la región S, digamos en ractángulos R_{ij}^S, las imágenes $\mathbf{f}(R_{ij}^S) = R_{ij}^K$ proporcionan una partición de la superficie K. En cada rectángulo R_{ij}^S tomemos un punto $(u_{ij}, v_{ij}) = \mathbf{p}_{ij}$, y consideremos la imagen de éste, $\mathbf{f}(\mathbf{p}_{ij}) = (x_{ij}, y_{ij}, z_{ij}) = \mathbf{q}_{ij}$, el cual se encontrará entonces en el "rectángulo" $R_{ij}^K = \mathbf{f}(R_{ij}^S) \in K$. Una estimación aproximada del flujo del campo $\mathbf{F}$ a través del rectángulo R_{ij}^K se obtiene al multiplicar la componente $F^{(2)}(\mathbf{q}_{ij})$ por el área del rectángulo correspondiente R_{ij}^K.

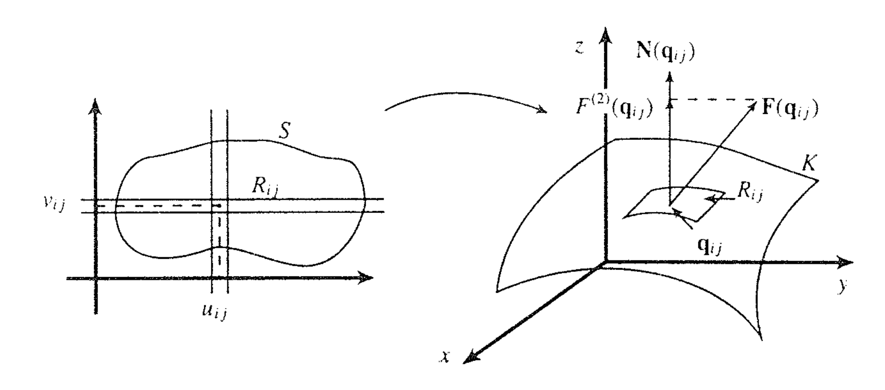


Figura 2. Las particiones de S y K y los vectores $\mathbf{N}(q_{ij})_{ij}$ y $\mathbf{F}(q_{ij})$

Es decir

$$\begin{aligned}\text{Flujo a través de } R_{ij}^K &\approx F^{(2)}(\mathbf{q}_{ij}) \text{ área de } (R_{ij}^K)\\ &= (\mathbf{F}(\mathbf{q}_{ij})\cdot\mathbf{N}(\mathbf{q}_{ij})) \text{ área de } R_{ij}^K) \approx (\mathbf{F}(\mathbf{q}_{ij})\cdot\mathbf{N}(\mathbf{q}_{ij}))\|\mathbf{N}(\mathbf{p}_{ij})\| \text{ área de } R_{ij}^S\\ &= (\mathbf{F}(\mathbf{f}(\mathbf{p}_{ij}))\cdot\mathbf{N}(\mathbf{f}(\mathbf{p}_{ij})))\|\mathbf{N}_\mathbf{f}(\mathbf{p}_{ij})\| \text{ área de } R_{ij}^S\end{aligned}$$

(donde el Área de $R_{ij}^K \approx \|\mathbf{N}_\mathbf{f}(\mathbf{p}_{ij})\|$ Área de R_{ij}^S, relación establecida en la sección 6 del capítulo 8).

Nótese que al sumar estos flujos sobre todos los rectángulos R_{ij}^K obtendremos una estimación del flujo a través de toda la superficie K. Es decir

$$\begin{aligned}\text{Flujo a través de } K &\approx \sum_{i,j} \text{flujo a través de } R_{ij}^K\\ &= \sum_{i,j}(\mathbf{F}(\mathbf{f}(\mathbf{p}_{ij}))\cdot\mathbf{N}(\mathbf{f}(\mathbf{p}_{ij})))\|\mathbf{N}_\mathbf{f}(\mathbf{p}_{ij})\| \text{ área de } R_{ij}^S\end{aligned}$$

Esta expresión no es más que una suma de Riemann de la función $\Phi: S \to \mathbb{R}$,

$$\Phi(u, v) = (\mathbf{F}(\mathbf{f}(u, v))\cdot\mathbf{N}(\mathbf{f}(u, v)))\|\mathbf{N}_\mathbf{f}(u, v)\|$$

sobre la región $S \subset \mathbb{R}^2$, de modo que si tomamos el límite cuando la cantidad de rectángulos crece al infinito (y sus áreas tienden a cero), obtenemos la integral doble de la función Φ sobre S. Es decir

$$\begin{aligned}\text{Flujo a través de } K &= \lim \sum_{i,j} \Phi(p_{ij}) \text{ área de } R_{ij}^S\\ &= \iint_S (\mathbf{F}(\mathbf{f}(u, v))\cdot\mathbf{N}(\mathbf{f}(u, v)))\|\mathbf{N}_\mathbf{f}(u, v)\| du\, dv\end{aligned}$$

Nótese que la integral doble anterior no es más que la integral de superficie de la función $\psi: K \to \mathbb{R}$, $\psi(x, y, z) = \mathbf{F}(x, y, z)\cdot\mathbf{N}(x, y, z)$ sobre la superficie K parametrizada por $\mathbf{f}$ (es decir, es una integral de superficie de las estudiadas en la sección anterior). Esta función ψ nos da en cada punto $(x, y, z) \in K$,

la componente del campo **F** en la dirección del campo de vectores normales a K. O sea, $\psi(x, y, z)$ nos dice "cuánto atraviesa de **F** a K en (x, y, z)". Es natural, que si "sumamos" (de manera continua) estas cantidades sobre todos los puntos $(x, y, z) \in K$, al resultado le llamemos "flujo de **F** a través de K". Esta es justamente la definición que ahora damos.

Definición. Sea $K = \mathbf{f}(S)$ una superficie simple parametrizada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, la cual proporciona una orientación que coincide con la del campo continuo de vectores normales $\mathbf{N}: K \to \mathbb{R}^3$. Sea $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ un campo continuo definido en el abierto U de $\mathbb{R}^3$ que contiene a K. Se define la integral de superficie de **F** sobre K (llamada flujo de **F** a través de K), denotada por $\int_{K_\mathbf{f}} \mathbf{F}\, dA$ como

$$\iint_{K_f} \mathbf{F}\, dA = \iint_K \mathbf{F} \cdot \mathbf{N}\, dA = \iint_S \mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N}(\mathbf{f}(u, v)) \|\mathbf{N_f}(u, v)\| du\, dv$$ ∎

Operaciones adicionales nos permiten presentar la integral de la definición anterior de una manera más simple. En efecto, puesto que el campo de vectores normales a K, $\mathbf{N}: K \to \mathbb{R}^3$ se puede presentar como

$$\mathbf{N}(\mathbf{q}) = \frac{\mathbf{N_f}(u, v)}{\|\mathbf{N_f}(u, v)\|}, \qquad \mathbf{q} = \mathbf{f}(u, v)$$

se tiene que

$$\begin{aligned}\mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N}(\mathbf{f}(u, v))\|\mathbf{N_f}(u, v)\| &= \mathbf{F}(\mathbf{f}(u, v)) \cdot \frac{\mathbf{N_f}(u, v)}{\|\mathbf{N_f}(u, v)\|} \|\mathbf{N_f}(u, v)\| \\ &= \mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N_f}(u, v) = \mathbf{F}(\mathbf{f}(u, v)) \cdot \left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right)\end{aligned}$$

Entonces

$$\iint_{K_\mathbf{f}} \mathbf{F}\, dA = \iint_S \mathbf{F}(\mathbf{f}(u, v)) \cdot \left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right) du\, dv$$

Antes de ver algunos ejemplos, tenga presente que la integral anterior es invariante por reparametrizaciones *que no cambian la orientación de la superficie K*. Si tomamos una reparametrización de K que cambie su orientación, esto sí se reflejará en un cambio de signo de la integral. En efecto, si $\mathbf{g} = \mathbf{f} \circ \varphi: S' \subset \mathbb{R}^2 \to \mathbb{R}^3$ es una reparametrización de **f**, en la que $\varphi: S' \to S$, $\varphi = (\varphi_1, \varphi_2)$, es una biyección de clase $\mathscr{C}^1$ con determinante jacobiano $\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)}$ nunca nulo en S', tenemos que, usando de nuevo la fórmula Â de la sección 2 del capítulo 8,

$$\begin{aligned}\iint_{K_\mathbf{f}} \mathbf{F}\, dA &= \iint_S \mathbf{F}(\mathbf{f}(u, v)) \cdot \left(\frac{\partial \mathbf{f}}{\partial u}(u, v) \times \frac{\partial \mathbf{f}}{\partial v}(u, v)\right) du\, dv \xrightarrow[(u, v) = \varphi(s, t)]{\text{cambio de variables}} \\ &= \iint_{S'} \mathbf{F}(\mathbf{f}(\varphi(s, t))) \cdot \left(\frac{\partial \mathbf{f}}{\partial u}(\varphi(s, t)) \times \frac{\partial \mathbf{f}}{\partial v}(\varphi(s, t))\right) \left|\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)}\right| ds\, dt \\ &= \iint_{S'} \mathbf{F}(\mathbf{g}(s, t)) \cdot \left(\frac{\partial \mathbf{g}}{\partial s}(s, t) \times \frac{\partial \mathbf{g}}{\partial t}(s, t)\right) \frac{\left|\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)}\right|}{\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)}} ds\, dt\end{aligned}$$

$$= \operatorname{sgn}\left(\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)}\right) \iint_{S'} \mathbf{F}(\mathbf{g}(s,t)) \cdot \left(\frac{\partial \mathbf{g}}{\partial s}(s,t) \times \frac{\partial \mathbf{g}}{\partial t}(s,t)\right) ds\, dt$$

$$= \operatorname{sgn}\left(\frac{\partial(\varphi_1, \varphi_2)}{\partial(s,t)}\right) \int_{K_{\mathbf{g}}} \mathbf{F}\, dA$$

Así

$$\iint_{K_{\mathbf{f}}} \mathbf{F}\, dA = \pm \int_{K_{\mathbf{g}}} \mathbf{F}\, dA$$

donde los signos $-$ ó $+$ corresponden a los casos en que la reparametrización $\mathbf{g}$ de $\mathbf{f}$ invierta o no la orientación de la superficie $K = \mathbf{f}(S) = \mathbf{g}(S')$, respectivamente. Así pues, estas nuevas integrales de superficie, en contraste con las estudiadas en la sección anterior, cambian de signo al cambiar la orientación del objeto matemático sobre el que se está integrando (la superficie K). Recuerde que éste es el mismo fenómeno que ocurrió con las integrales de línea respecto de la longitud de arco (sección 7, capítulo 7), las cuales no cambiaban de signo al cambiar de orientación el camino sobre el que se estaba integrando, y las integrales de línea de campos vectoriales (sección 3, capítulo 7), en las que un cambio de orientación del camino $\boldsymbol{\lambda}$ sobre el que se integraba, se reflejaba en un cambio de signo en la integral.

NOTA: La notación $\int_{K_{\mathbf{f}}} \mathbf{F}\, dA$ que hemos usado para estas integrales, sugiere que la superficie K tiene una orientación determinada por la función $\mathbf{f}$ que la parametriza. Así, tal notación invita a pensar que el objeto matemático sobre el que se está integrando es la superficie K *junto con la orientación proporcionada por la parametrización* $\mathbf{f}$.

Para los cálculos de flujos de campos a través de una superficie se tiene que proporcionar entonces (además, por supuesto, del campo $\mathbf{F}$ y la superficie K) la *orientación* considerada de la superficie K. Dada ésta, digamos que en forma de un campo de vectores continuo $\tilde{\mathbf{N}}: K \to \mathbb{R}^3$, debemos procurar una parametrización $\mathbf{f}$ que "coincida" con la orientación dada. Más aún, $\mathbf{f}$ debe ser tal que $\frac{\partial \mathbf{f}}{\partial u}(u, v) \times \frac{\partial \mathbf{f}}{\partial v}(u, v) = c\tilde{\mathbf{N}}(\mathbf{f}(u, v))$ para cada $(u, v) \in S$, donde c es una *constante positiva*.

Ejemplo 1. Sea K la superficie constituida por la porción del plano $z = \alpha x + \beta y + \gamma$, en la región $S = [a, b] \times [c, d]$, con la orientación dada por sus vectores normales $(-\alpha, -\beta, 1)$ en cada uno de sus puntos. La función $\mathbf{f}: S \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \alpha u + \beta v + \gamma)$ es una parametrización que proporciona la orientación dada, pues

$$\mathbf{N}_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = (1, 0, \alpha) \times (0, 1, \beta) = (-\alpha, -\beta, 1)$$

Consideremos el campo $\mathbf{f}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (0, 1, 0)$. Este es un campo constante que fluye en paralelo al eje y en su dirección positiva. El flujo de $\mathbf{F}$ a través de K es

$$\int_{K_{\mathbf{f}}} \mathbf{F}\, dA = \iint_S \mathbf{F}(\mathbf{f}(u, v)) \cdot \left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right) du\, dv$$

$$= \iint_S (0, 1, 0) \cdot (-\alpha, -\beta, 1) du\, dv = \int_a^b du \int_c^d -\beta\, dv = -\beta(b - a)(d - c)$$

El resultado depende sólo de las dimensiones de la región S y de la coordenada β del vector normal al plano. Nótese que

1. Si $\beta = 0$ el flujo de $\mathbf{F}$ a través del plano es 0. En este caso el flujo *es paralelo* al plano (es decir, $\mathbf{F}(\mathbf{q})$ es perpendicular a $(-\alpha, -\beta, 1)$, pues $(0, 1, 0) \cdot (-\alpha, -\beta, 1) = -\beta = 0$), por lo que los vectores $\mathbf{F}(\mathbf{q})$ *no atraviesan la superficie* K.

2. Si $\beta > 0$, el flujo de $\mathbf{F}$ a través de K es negativo. En este caso el vector $(-\alpha, -\beta, 1)$ apunta "en dirección contraria" a (la componente normal de) $\mathbf{F}(\mathbf{q})$, $\mathbf{q} \in K$.

3. Si $\beta < 0$, el flujo de $\mathbf{F}$ a través de K es positivo. En este caso el vector $(-\alpha, -\beta, 1)$ apunta "en la misma dirección" a (la componente normal de) $\mathbf{F}(\mathbf{q})$, $\mathbf{q} \in K$.

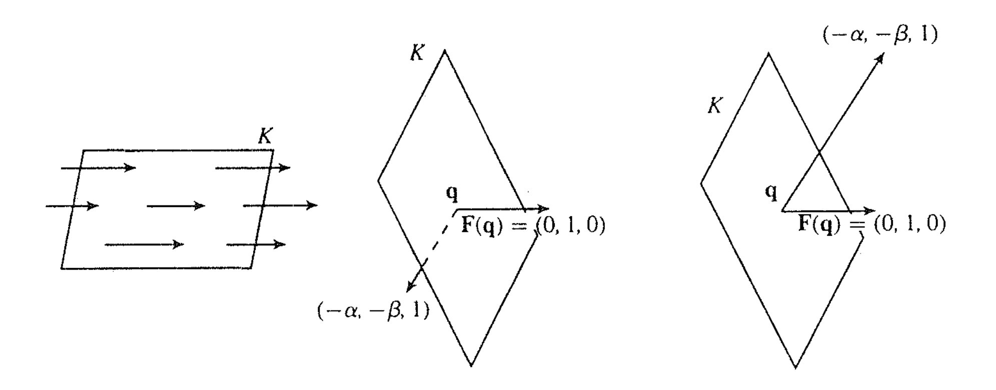


Figura 3. Los tres casos del ejemplo 1. ■

Ejemplo 2. Sea K la porción de la superficie parabólica $z = u^2 + v^2$ sobre la región $S = \{(u, v)|u^2 + v^2 \leq 1\}$, con sus normales apuntando "hacia afuera" de ella. Esta superficie ha aparecido en repetidas ocasiones en el capítulo anterior y en éste, y hemos manejado ya varias veces la función $\mathbf{f}: S \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, u^2 + v^2)$ como una parametrización de ella. Esta función proporciona una orientación de K dada por

$$\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = (1, 0, 2u) \times (0, 1, 2v) = (-2u, -2v, 1)$$

Nótese que esta orientación es la opuesta a la proporcionada en el ejemplo, pues por ejemplo, en el punto $\mathbf{f}(0, 0) = (0, 0, 0) \in K$, se tiene $\mathbf{N_f}(0, 0) = (0, 0, 1)$, y este vector apunta "hacia adentro" del paraboloide. Nos encontramos entonces con el problema de procurar una reparametrización de $\mathbf{f}$ que produzca la orientación opuesta. La manera de hacerlo es componiendo $\mathbf{f}$ con una función $\varphi: S' \to S$ biyectiva, cuyo determinante jacobiano sea negativo. Digamos, por ejemplo, que la función $\varphi = (\varphi_1, \varphi_2)$ tenga por jacobiano a

$$\frac{\partial(\varphi_1, \varphi_2)}{\partial(s, t)} = \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1$$

Es claro que una función semejante es $\varphi(s, t) = (t, s)$, y el efecto que producirá al componer $\mathbf{f}$ con φ será simplemente "invertir" las variables u y v (obsérvese que en este caso $S' = S$). Así pues, la

función $\mathbf{g}: S \to \mathbb{R}^3$, $\mathbf{g}(u, v) = \mathbf{f}(v, u) = (v, u, u^2 + v^2)$ parametriza también a K y ahora la orientación que ésta proporciona resulta de

$$\mathbf{N}_{\mathbf{g}}(u, v) = \frac{\partial \mathbf{g}}{\partial u} \times \frac{\partial \mathbf{g}}{\partial v} = (0, 1, 2u) \times (1, 0, 2v) = (2u, 2v, -1)$$

la cual es la orientación que aparece en el ejemplo. Sea $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ el campo $\mathbf{F}(x, y, z) = (x, y, z - 1)$. El flujo de $\mathbf{F}$ a través de K es

$$\iint_{K_g} \mathbf{F}\, dA = \iint_S \mathbf{F}(\mathbf{g}(u, v)) \cdot \left(\frac{\partial \mathbf{g}}{\partial u} \times \frac{\partial \mathbf{g}}{\partial v}\right) du\, dv$$

$$\iint_S (v, u, u^2 + v^2 - 1) \cdot (2u, 2v, -1) du\, dv = \iint_S (4uv + 1 - u^2 - v^2) du\, dv \overset{\text{polares}}{=}$$

$$= \int_0^{2\pi} \int_0^1 (4r^2 \cos\theta \operatorname{sen}\theta + 1 - r^2) r\, dr\, d\theta = \frac{\pi}{2} \quad \blacksquare$$

La definición de flujo de un campo $\mathbf{F}$ a través de una superficie simple K, funciona también para superficies más generales en las que tenga siempre sentido hablar del campo continuo $\mathbf{N}_{\mathbf{f}}(u, v)$ proporcionado por una parametrización de K, con el cual la superficie queda orientada.

Ejemplo 3. Sea K la esfera $x^2 + y^2 + z^2 = c^2$ con sus normales apuntando hacia afuera de ella. Como se ha manejado ya en varias ocasiones, una parametrización de K está dada por las coordenadas esféricas $\mathbf{f}: [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$

$$\mathbf{f}(u, v) = (c \cos u \operatorname{sen} v, c \operatorname{sen} u \operatorname{sen} v, c \cos v)$$

en cuyo caso

$$\mathbf{N}_{\mathbf{f}}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = -c^2(\cos u \operatorname{sen}^2 v, \operatorname{sen} u \operatorname{sen}^2 v, \operatorname{sen} v \cos v)$$

(ver ejemplo 8 de la sección 4 del capítulo 8). Nuevamente vemos que esta parametrización proporciona una orientación contraria a la dada, pues, por ejemplo, en el punto $\mathbf{q} = \mathbf{f}\left(0, \frac{\pi}{2}\right) = (c, 0, 0) \in K$ se tiene

$$\mathbf{N}_{\mathbf{f}}\left(0, \frac{\pi}{2}\right) = -c^2(1, 0, 0)$$

el cual es un vector que apunta hacia "adentro" de la esfera en $(c, 0, 0)$.

De nuevo, como en el ejemplo 2, podemos considerar la parametrización $\mathbf{g}: S' = [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$,

$$\mathbf{g}(u, v) = \mathbf{f}(v, u) = (c \cos v \operatorname{sen} u, c \operatorname{sen} v \operatorname{sen} u, c \cos u)$$

para la cual

$$\mathbf{N}_{\mathbf{g}}(u, v) = \frac{\partial \mathbf{g}}{\partial u} \times \frac{\partial \mathbf{g}}{\partial v} = c^2(\cos v \operatorname{sen}^2 u, \operatorname{sen} v \operatorname{sen}^2 u, \operatorname{sen} u \cos u)$$

Ahora se tiene en el punto $\mathbf{q} = \mathbf{f}\left(0, \frac{\pi}{2}\right) = \mathbf{g}\left(\frac{\pi}{2}, 0\right) = (c, 0, 0) \in K$ el vector normal

$$\mathbf{N}_{\mathbf{g}}(u, v) = c^2(1, 0, 0)$$

el cual apunta hacia el exterior de la esfera.

Consideremos el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ dado por

$$\mathbf{F}(x, y, z) = (x, y, z)$$

y calculemos el flujo de $\mathbf{F}$ a través de K. Se tiene

$$\iint_{K_g} \mathbf{F}\, dA = \iint_{S'} \mathbf{F}(\mathbf{g}(u, v)) \cdot \left(\frac{\partial \mathbf{g}}{\partial u} \times \frac{\partial \mathbf{g}}{\partial v}\right) du\, dv$$

$$= \iint_{S'} (c \cos v \operatorname{sen} u, c \operatorname{sen} v \operatorname{sen} u, c \cos u) \cdot (c^2 \cos v \operatorname{sen}^2 u, c^2 \operatorname{sen} v \operatorname{sen}^2 u, c^2 \operatorname{sen} u \cos u) du\, dv$$

$$= \int_0^{2\pi} dv \int_0^{\pi} c^3 \operatorname{sen} u\, du = 4\pi c^3$$

Otra manera de abordar este problema es la siguiente: un campo de vectores normales a la esfera $x^2 + y^2 + z^2 = c^2$ que apunta hacia el exterior de ésta es $\mathbf{N}: K \to \mathbb{R}^3$, $\mathbf{N}(x, y, z) = \frac{1}{\|(x, y, z)\|}(x, y, z) = \frac{1}{c}(x, y, z)$, de modo que

$$\mathbf{F}(x, y, z) \cdot \mathbf{N}(x, y, z) = (x, y, z) \cdot \frac{1}{c}(x, y, z) = \frac{1}{c}(x^2 + y^2 + z^2)$$

Al evaluar este producto punto en los puntos de la esfera, en los que $x^2 + y^2 + z^2 = c^2$, se obtiene

$$\mathbf{F} \cdot \mathbf{N} = \frac{1}{c}(c^2) = c$$

como $\|\mathbf{N}_g(u, v)\| = c^2 \operatorname{sen} u$, se tiene

$$\iint_{K_g} \mathbf{F}\, dA = \iint_{S'} \mathbf{F} \cdot \mathbf{N} \|\mathbf{N}_g(u, v)\| du\, dv = \iint_{S'} c(c^2 \operatorname{sen} u) du\, dv$$

$$= \int_0^{2\pi} dv \int_0^{\pi} c^3 \operatorname{sen} u du = 4\pi c^3$$

■

Ejemplo 4. Consideremos el campo $\mathbf{F}: \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}^3$ dado por

$$\mathbf{F}(x, y, z) = \frac{\alpha}{(x^2 + y^2 + z^2)^{3/2}}(x, y, z)$$

donde α es una constante. Este es un campo "muy parecido" al del ejemplo anterior: ambos son campos *radiales*. Su única diferencia es la magnitud que tienen en cada punto (x, y, z) de $\mathbb{R}^3 - \{(0, 0, 0)\}$. Observe que la magnitud del campo considerado es

$$\|\mathbf{F}(x, y, z)\| = \frac{\alpha}{\|(x, y, z)\|^2}$$

Es decir, su magnitud en el punto $(x, y, z) \in \mathbb{R}^3 - \{(0, 0, 0)\}$ es proporcional al inverso del cuadrado de la distancia de (x, y, z) al origen. Este tipo de campos son muy importantes en ciertas áreas de la física.

Ejemplos de ellos son el campo gravitacional y el campo eléctrico (ver sección 8 del capítulo 7). Calculemos el flujo de este campo a través de la esfera $x^2 + y^2 + z^2 = c^2$ con sus vectores normales apuntando hacia el exterior de ésta. Usando la misma idea de la parte final del ejemplo anterior, tenemos que un campo de vectores normales a la esfera es $\mathbf{N}\colon K \to \mathbb{R}^3$, $\mathbf{N}(x, y, z) = \frac{1}{c}(x, y, z)$, de modo que

$$\mathbf{F}(x, y, z) \cdot \mathbf{N}(x, y, z) = \frac{\alpha/c}{(x^2 + y^2 + z^2)^{3/2}}(x^2 + y^2 + z^2)$$
$$= \frac{\alpha/c}{(x^2 + y^2 + z^2)^{1/2}} = \frac{\alpha/c}{\|(x, y, z)\|}$$

En los puntos (x, y, z) de la esfera se tiene $\|(x, y, z)\| = c$, de modo que el producto punto anterior se ve como

$$\mathbf{F} \cdot \mathbf{N} = \frac{\alpha/c}{c} = \frac{\alpha}{c^2}$$

Puesto que $\|\mathbf{N}_{\mathbf{g}}(u, v)\| = c^2 \operatorname{sen} u$, donde $\mathbf{g}\colon S' = [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$, $\mathbf{g}(u, v) = (c \cos v \operatorname{sen} u$, $\operatorname{sen} v \operatorname{sen} u, c \cos u)$, es la función que parametriza a la esfera (junto con la orientación establecida), se tiene que el flujo de $\mathbf{F}$ a través de la esfera es

$$\iint_{K_{\mathbf{g}}} \mathbf{F}\, dA = \iint_{S'} (\mathbf{F} \cdot \mathbf{N})(\mathbf{g}(u, v))\|\mathbf{N}_{\mathbf{g}}(u, v)\|\, du\, dv$$
$$= \iint_{S'} \frac{\alpha}{c^2}(c^2 \operatorname{sen} u) du\, dv = \alpha \int_0^{2\pi} dv \int_0^{\pi} \operatorname{sen} u du = 4\pi\alpha$$

Un hecho notable en este resultado es que el flujo del campo $\mathbf{F}$ a través de la esfera $x^2 + y^2 + z^2 = c^2$, *no depende del radio de la esfera*. Esta situación es un caso especial de un resultado que se desprende del Teorema de la Divergencia que se estudiará en la próxima sección. ■

Ejemplo 5. Sea K el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

con sus vectores normales apuntando hacia su exterior. Con la parametrización de la esfera del ejemplo 3 no será difícil aceptar que una parametrización de K que produce la orientación dada es $\mathbf{g}\colon S = [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$

$$\mathbf{g}(u, v) = (a \cos v \operatorname{sen} u, b \operatorname{sen} v \operatorname{sen} u, c \cos u)$$

En este caso se tiene

$$\mathbf{N}_{\mathbf{g}}(u, v) = \frac{\partial \mathbf{g}}{\partial u} \times \frac{\partial \mathbf{g}}{\partial v} = (bc \cos v \operatorname{sen}^2 u, ac \operatorname{sen} v \operatorname{sen}^2 u, ab \operatorname{sen} u \cos u)$$

Consideremos el mismo campo $\mathbf{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ del ejemplo 3, $\mathbf{F}(x, y, z) = (x, y, z)$. El flujo de $\mathbf{F}$ a través de K es

$$\iint_{K_g} \mathbf{F}\, dA = \iint_S \mathbf{F}(\mathbf{g}(u,v)) \cdot \mathbf{N_g}(u,v) du\, dv$$

$$= \iint_S (a\cos v \operatorname{sen} u, b \operatorname{sen} v \operatorname{sen} u, c\cos u) \cdot (bc\cos v \operatorname{sen}^2 u, ac \operatorname{sen} v \operatorname{sen}^2 u, ab \operatorname{sen} u \cos u)\, du\, dv$$

$$= \int_0^{2\pi} dv \int_0^{\pi} abc \operatorname{sen} u\, du = 4\pi abc$$

En particular, si $a = b = c$ se obtiene el resultado del ejemplo 3. ■

Ejemplo 6. Consideremos nuevamente el elipsoide K del ejemplo anterior parametrizado por $\mathbf{g}: S = [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$, $\mathbf{g}(u, v) = (a \cos v \operatorname{sen} u, b \operatorname{sen} v \operatorname{sen} u, c \cos u)$. Sea $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ un campo constante, digamos $\mathbf{F}(x, y, z) = (\alpha, \beta, \gamma)$. El flujo de $\mathbf{F}$ a través de K es

$$\iint_{K_g} \mathbf{F}\, dA = \iint_S \mathbf{F}(\mathbf{g}(u,v)) \cdot \mathbf{N_g}(u,v) du\, dv$$

$$= \int_S\int (\alpha, \beta, \gamma) \cdot (bc\cos v \operatorname{sen}^2 u, ac \operatorname{sen} v \operatorname{sen}^2 u, ab \operatorname{sen} u \cos u) du\, dv$$

$$= \int_0^{2\pi} dv \int_0^{\pi} (\alpha bc\cos v \operatorname{sen}^2 u + \beta ac \operatorname{sen} v \operatorname{sen}^2 u + \gamma ab \operatorname{sen} u \cos u) du = 0$$

Este resultado nos dice que el flujo NETO de $\mathbf{F}$ a través de K es cero: no importa el campo $\mathbf{F}$ (con tal de que sea constante), lo que "entra de $\mathbf{F}$ a través de K" es igual a lo que "sale de $\mathbf{F}$ a través de K". ■

Ejemplo 7. Sea K el cubo formado por los 6 planos $x = \pm a$, $y = \pm a$, $z = \pm a$ (donde $a > 0$), con sus vectores normales apuntando hacia el exterior de K. Se trata entonces de una superficie formada por "la unión" de 6 superficies simples. El flujo de un campo $\mathbf{F}$ a través de K se puede calcular sumando los flujos a través de cada una de las 6 caras de K. Consideremos, por ejemplo, un campo constante $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (\alpha, \beta, \gamma)$, donde α, β y γ son positivos (es decir, las flechas de $\mathbf{F}$ apuntan hacia el primer octante). Para la cara correspondiente al plano $x = a$, se tiene que una parametrización es $\mathbf{f}: [-a, a] \times [-a, a] \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (a, u, v)$. Esta parametrización orienta la cara correspondiente hacia el exterior (como el ejemplo establece que debe ocurrir), pues $\mathbf{N_f}(u, v) = (1, 0, 0)$. Se tiene entonces que el flujo de $\mathbf{F}$ a través de esta cara del cubo es

$$\iint_{K_f} \mathbf{F}\, dA = \iint_{[-a,a]\times[-a,a]} \mathbf{F}(\mathbf{f}(u,v)) \cdot \mathbf{N_f}(u,v) du\, dv$$

$$\int_{-a}^{a}\int_{-a}^{a} (\alpha, \beta, \gamma) \cdot (1, 0, 0) du\, dv = \int_{-a}^{a}\int_{-a}^{a} \alpha du\, dv = 4a\alpha$$

Los cálculos con las cinco caras restantes son análogos. Se resumen éstos en el cuadro siguiente

Cara	Parametrización $\mathbf{f}: [-a, a] \times [-a, a] \to \mathbb{R}^3$	$\mathbf{N_f}(u, v)$	$\mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N_f}(u, v)$	flujo de **F** a través de la cara
$x = a$	$\mathbf{f}(u, v) = (a, u, v)$	$(1, 0, 0)$	α	$4a\alpha$
$x = -a$	$\mathbf{f}(u, v) = (-a, u, v)$	$(-1, 0, 0)$	$-\alpha$	$-4a\alpha$
$y = a$	$\mathbf{f}(u, v) = (v, a, u)$	$(0, 1, 0)$	β	$4a\beta$
$y = -a$	$\mathbf{f}(u, v) = (u, -a, v)$	$(0, -1, 0)$	$-\beta$	$-4a\beta$
$z = a$	$\mathbf{f}(u, v) = (u, v, a)$	$(0, 0, 1)$	γ	$4a\gamma$
$z = -a$	$\mathbf{f}(u, v) = (v, u, -a)$	$(0, 0, -1)$	$-\gamma$	$-4a\gamma$

El flujo de **F** a través de K es entonces

$$(4a\alpha) + (-4a\alpha) + (4a\beta) + (-4a\beta) + (4a\gamma) + (-4a\gamma) = 0$$

Observe los signos de los flujos de **F** a través de las caras del cubo: lo que entra en una de ellas es lo mismo que lo que sale en la cara opuesta (como es natural esperarse pues el campo **F** es constante). Así pues, al sumar todos estos flujos se obtiene un flujo neto de **F** a través del cubo igual a cero. ∎

Terminamos esta sección con un comentario sobre otra notación usada para las integrales de campos vectoriales, la cual dejará ver (todavía más) el parentezco que éstas tienen (como las generalizaciones a $\mathbb{R}^3$) con las integrales de línea estudiadas en la sección 3 del capítulo 7. Recuerde: si $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (F_1, F_2)$, es un campo continuo definido en el conjunto abierto U de $\mathbb{R}^2$ y $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (x(t), y(t))$, es un camino de clase $\mathscr{C}^1$ tal que $\boldsymbol{\lambda}([a, b]) \subset U$, entonces la integral de línea de **F** a lo largo de $\boldsymbol{\lambda}$ se escribe como

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}} F_1(x, y)\,dx + F_2(x, y)\,dy$$

lo cual quería decir (se calculaba como)

$$\int_a^b (F_1(\boldsymbol{\lambda}(t))x'(t) + F_2(\boldsymbol{\lambda}(t))y'(t))dt$$

Para las integrales de superficie estudiadas en esta sección se tiene un campo continuo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ definido en el conjunto abierto U de $\mathbb{R}^3$ y $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ la parametrización de una superficie simple $K = \mathbf{f}(S)$, con $K \subset U$ (la cual proporciona una orientación determinada para K). Se definió la integral de superficie de **F** sobre K como

$$\iint_{K_f} \mathbf{F}\,dA = \iint_K \mathbf{F} \cdot \mathbf{N}\,dA = \iint_S \mathbf{F}(\mathbf{f}(u, v)) \cdot \left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right) du\,dv$$

Poniendo $\mathbf{F} = (F_1, F_2, F_3)$ y $\mathbf{f} = (f_1, f_2, f_3)$, donde $F_1, F_2, F_3: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ son funciones continuas y $f_1, f_2, f_3: S \subset \mathbb{R}^2 \to \mathbb{R}$ son funciones de clase $\mathscr{C}^1$, podemos escribir

$$\left(\frac{\partial \mathbf{f}}{\partial u}\right) \times \left(\frac{\partial \mathbf{f}}{\partial v}\right) = \left(\frac{\partial f_1}{\partial u}, \frac{\partial f_2}{\partial u}, \frac{\partial f_3}{\partial u}\right) \times \left(\frac{\partial f_1}{\partial v}, \frac{\partial f_2}{\partial v}, \frac{\partial f_3}{\partial v}\right)$$
$$= \left(\frac{\partial(f_2, f_3)}{\partial(u, v)}, \frac{\partial(f_1, f_3)}{\partial(u, v)}, \frac{\partial(f_1, f_2)}{\partial(u, v)}\right)$$

(ver sección 1 del capítulo 8), de modo que, si llamamos $x = f_1(u, v)$, $y = f_2(u, v)$, $z = f_3(u, v)$, tenemos

$$\left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right) = \left(\frac{\partial(y, z)}{\partial(u, v)}, \frac{\partial(z, x)}{\partial(u, v)}, \frac{\partial(x, y)}{\partial(u, v)}\right)$$

Si se desea, se puede identificar "el producto de los diferenciales" de funciones de dos variables, digamos $\boldsymbol{\tau} = \boldsymbol{\tau}(u, v)$, $\boldsymbol{\eta} = \boldsymbol{\eta}(u, v)$, como

$$d\boldsymbol{\tau} d\boldsymbol{\eta} = \frac{\partial(\boldsymbol{\tau}, \boldsymbol{\eta})}{\partial(u, v)}\, du\, dv$$

de modo que

$$\left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right) = (dy\, dz, dz\, dx, dx\, dy)$$

Se tiene entonces que

$$\begin{aligned}\mathbf{F}(\mathbf{f}(u, v)) \cdot \left(\frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}\right) du\, dv &= (F_1(\mathbf{f}(u, v)), F_2(\mathbf{f}(u, v)), F_3(\mathbf{f}(u, v))) \cdot (dy\, dz, dz\, dx, dx\, dy) \\ &= F_1(\mathbf{f}(u, v)) dy\, dz + F_2(\mathbf{f}(u, v)) dz\, dx + F_3(\mathbf{f}(u, v)) dx\, dy\end{aligned}$$

Y al fin

$$\iint_{K_{\mathbf{f}}} \mathbf{F}\, dA = \iint_{K_{\mathbf{f}}} F_1 dy\, dz + F_2 dz\, dx + F_3 dx\, dy$$

la cual es una notación que recuerda la ya antes mencionada de las integrales de línea. Veámoslas juntas:

campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (F_1, F_2)$ camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (x(t), y(t))$ $\boldsymbol{\lambda}([a, b]) \subset U$ $\int_{\boldsymbol{\lambda}} \mathbf{F}\, d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}} F_1\, dx + F_2\, dy$ $= \int_a^b \left(F_1(\boldsymbol{\lambda}(t))x'(t) + F_2(\boldsymbol{\lambda}(t))y'(t)\right) dt$	campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$ superficie $\mathbf{f}: S \subseteq \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (x(u, v), y(u, v), z(u, v))$ $\mathbf{f}(S) \subset U$ $\iint_{K_{\mathbf{f}}} \mathbf{F}\, dA = \iint_{K_{\mathbf{f}}} F_1\, dydz + F_2\, dzdx + F_3\, dxdy$ $= \iint_S \left(F_1(\mathbf{f}(u, v))\frac{\partial(y, z)}{\partial(u, v)}\right.$ $+ F_2(\mathbf{f}(u, v))\frac{\partial(z, x)}{\partial(u, v)}$ $\left.+ F_3(\mathbf{f}(u, v))\frac{\partial(x, y)}{\partial(u, v)}\right) du\, dv$

Estas "coincidencias" se verán en la óptica adecuada en el capítulo 10.

Ejercicios (Capítulo 9, Sección 2)

1. Sea $K = \mathbf{f}(S)$ una superficie simple, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ (de la que hereda la orientación). Sean $F_1, F_2: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ dos campos continuos definidos en el conjunto abierto U que contiene a K. Demuestre que

$$\iint_{K_f} (F_1 + cF_2)\, dA = \iint_{K_f} F_1\, dA + c \iint_{K_f} F_2\, dA$$

donde c es un número real.

En los ejercicios 2—12, calcule el flujo del campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ a través de la superficie K.

2. $\mathbf{F}(x, y, z) = (1, 0, 0)$, a través de la porción del plano yz, correspondiente a $-1 \leq y \leq 1$, $-1 \leq z \leq 1$, con sus normales apuntando en la dirección del vector $(1, 0, 0)$.

3. $\mathbf{F}(x, y, z) = (0, xyz, x^2y^2z^2)$, a través de la porción del plano yz, correspondiente a $-1 \leq y \leq 1$, $-1 \leq z \leq 1$, con sus normales apuntando en la dirección del vector $(-1, 0, 0)$.

4. $\mathbf{F}(x, y, z) = (x, y, z)$, a través de la porción del plano $x + y + z = 1$ en el primer octante, con sus normales apuntando hacia el "exterior" (con la tercera coordenada positiva).

5. $\mathbf{F}(x, y, z) = (x, y, z)$, a través de la superficie $|x| + |y| + |z| = 1$, con sus normales apuntando hacia su exterior. (Sugerencia: use el resultado del ejercicio anterior).

6. $\mathbf{F}(x, y, z) = (z^2, 0, 0)$, a través del semielipsoide superior $2x^2 + 3y^2 + z^2 = 6$, $z \geq 0$, con sus normales apuntando hacia su exterior.

7. $\mathbf{F}(x, y, z) = (0, x^2, 0)$, a través del semielipsoide superior $2x^2 + 3y^2 + z^2 = 6$, $z \geq 0$, con sus normales apuntando hacia su exterior.

8. $\mathbf{F}(x, y, z) = (0, 0, y^2)$, a través del semielipsoide superior $2x^2 + 3y^2 + z^2 = 6$, $z \geq 0$, con sus normales apuntando hacia su exterior.

9. $\mathbf{F}(x, y, z) = (3z^2, 4x^2, 5y^2)$, a través del semielipsoide superior $2x^2 + 3y^2 + z^2 = 6$, $z \geq 0$, con sus normales apuntando hacia su exterior. (Sugerencia: use los resultados de los tres ejercicios anteriores, así como el del ejercicio 1).

10. $\mathbf{F}(x, y, z) = (xz, xy, yz)$, a través del tetraedro formado por los planos coordenados y el plano $x + y + z = 1$, con sus normales apuntando hacia su exterior.

11. $\mathbf{F}(x, y, z) = (2xy, z, y)$, a través del cilindro $x^2 + y^2 = 1$, $-1 \leq z \leq 1$, con sus normales apuntando hacia su interior.

12. $\mathbf{F}(x, y, z) = (x, x + y, x + y + z)$, a través de la porción del plano $2x + 3y + 6z = 1$ que se encuentra en el primer octante, con sus normales apuntando hacia su "interior" (con su tercera coordenada negativa).

13. Demuestre que el flujo del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (ax, by, cz)$, donde a, b, c son números reales positivos, a través de la esfera con centro en el origen y radio $c > 0$, con sus normales apuntando al exterior, es igual a $(a + b + c)$ veces el volumen de la esfera.

14. Demuestre que el flujo del campo $\mathbf{F}$ del ejercicio anterior a través del cubo $K = [-1, 1] \times [-1, 1] \times [-1, 1]$, con sus normales apuntando hacia su exterior, es igual a $8(a+b+c) = (a+b+c)$ veces el volumen del cubo.

15. Demuestre que el flujo del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (y + z, x + z, x + y)$, a través del cubo $K = [-1, 1] \times [-1, 1] \times [-1, 1]$, con sus normales apuntando hacia su exterior, es igual a cero.

16. (Generalización del ejercicio anterior). Sea $f: \mathbb{R}^2 \to \mathbb{R}$ una función continua. Considere el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ resultante de $\mathbf{F}(x, y, z) = (f(y, z), f(x, z), f(x, y))$. Demuestre que el flujo del campo $\mathbf{F}$ a través del cubo $K = [-1, 1] \times [-1, 1] \times [-1, 1]$, con sus normales apuntando hacia su exterior, es igual a cero.

9.3 La divergencia de un campo vectorial (II):campos en $\mathbb{R}^n$

En esta sección se prosigue con la exposición de la sección 11 del capítulo 7, en la que se introdujo el concepto de divergencia para un campo vectorial en $\mathbb{R}^2$. En aquella sección se presentó a la divergencia de un campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ (diferenciable) como una medida del flujo del campo por unidad de área en cada punto $\mathbf{p} \in U$. Se dijo que esta idea podía ser generalizada para campos en $\mathbb{R}^3$, donde ahora la divergencia de un campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ será una medida del flujo del campo por unidad de volumen en cada punto $\mathbf{p} \in U$. Así pues, la idea desarrollada en el capítulo 7 se puede reproducir ahora con un campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, tomando, por ejemplo un "pequeño cubo" $K_{\mathbf{p}}$ que tenga al punto $\mathbf{p} \in U$ en su centro, calculando el flujo de $\mathbf{F}$ a través de $K_{\mathbf{p}}$ (con las integrales de superficie de la sección anterior), dividiendo entre el volumen de $K_{\mathbf{p}}$, y tomando el límite cuando el cubo $K_{\mathbf{p}}$ se "contrae" hacia el punto $\mathbf{p}$. El resultado al que se llegará es

$$\lim_{K_{\mathbf{p}} \to \mathbf{p}} \frac{\text{flujo de } \mathbf{F} \text{a través de } K_{\mathbf{p}}}{\text{volumen de } K_{\mathbf{p}}} = \frac{\partial F_1}{\partial x}(\mathbf{p}) + \frac{\partial F_2}{\partial y}(\mathbf{p}) + \frac{\partial F_3}{\partial z}(\mathbf{p})$$

donde $\mathbf{F} = (F_1, F_2, F_3)$. La expresión a la derecha en esta fórmula es lo que llamamos divergencia de $\mathbf{F}$ en $\mathbf{p}$. En general, recuerde que para un campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, diferenciable, la divergencia de $\mathbf{F}$ en $\mathbf{p} \in U$, denotada por div $\mathbf{F}(\mathbf{p})$, es

$$\operatorname{div} \mathbf{F}(\mathbf{p}) = \frac{\partial F_1}{\partial x_1}(\mathbf{p}) + \frac{\partial F_2}{\partial x_2}(\mathbf{p}) + \cdots + \frac{\partial F_n}{\partial x_n}(\mathbf{p})$$

Dejamos como ejercicio para el lector las operaciones arriba mencionadas para obtener así la expresión de la divergencia del campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ diferenciable, en un punto $\mathbf{p} \in U$. Nuestro interés fundamental en esta sección será establecer el análogo del teorema 7.11.1 para campos en $\mathbb{R}^3$. Es decir, establecer el teorema de la Divergencia para campos en $\mathbb{R}^3$. Veremos que con este resultado es posible obtener la "definición" de div $\mathbf{F}(\mathbf{p})$ como el límite del cociente del flujo del campo $\mathbf{F}$ a través de un cubo que circunda a $\mathbf{p}$ dividido entre el volumen del cubo, cuando éste se contrae hacia $\mathbf{p}$.

El Teorema de la Divergencia (para campos en $\mathbb{R}^3$) relaciona la integral de superficie del campo $\mathbf{F}$ sobre una superficie cerrada K (es decir, el flujo de $\mathbf{F}$ a través de K) con la integral triple de div $\mathbf{F}$ sobre la región Ω que encierra K. Las superficies que vamos a considerar serán fronteras de regiones

$\Omega \subset \mathbb{R}^3$ que se puede ver como regiones de los tres tipos considerados en el capítulo 6 (sección 6). Es decir, $\Omega \subset \mathbb{R}^3$ se puede describir como

$$\begin{aligned}\Omega &= \{(x, y, z)|(x, y) \in R_1, \varphi_1(x, y) \leq z \leq \varphi_2(x, y)\} \\ &= \{(x, y, z)|(x, z) \in R_2, \psi_1(x, z) \leq y \leq \psi_2(x, z)\} \\ &= \{(x, y, z)|(y, z) \in R_3, \nu_1(y, z) \leq x \leq \nu_2(y, z)\}\end{aligned}$$

para ciertas funciones (que necesitamos que sean de clase $\mathscr{C}^1$) φ_1, φ_2: $R_1 \subset \mathbb{R}^2 \to \mathbb{R}$ (de las variables xy), ψ_1, ψ_2: $R_2 \subset \mathbb{R}^2 \to \mathbb{R}$ (de las variables xz), ν_1, ν_2: $R_3 \subset \mathbb{R}^2 \to \mathbb{R}$ (de las variables yz), definidas en las regiones R_1, R_2, R_3 de los planos xy, xz y yz respectivamente. O sea que Ω se puede ver limitada "por arriba y por abajo, por gráficas de funciones $z = \varphi(x, y)$", "por la derecha y por la izquierda, por gráficas de funciones $y = \psi(x, z)$" y "por delante y por detrás, por gráficas de funciones $x = \nu(y, z)$".

La frontera de la región Ω es una superficie K en $\mathbb{R}^3$. Por ejemplo, viendo Ω como una región del tipo

$$\{(x, y, z)|(x, y) \in R_1, \varphi_1(x, y) \leq z \leq \varphi_2(x, y)\}$$

la frontera de Ω está constituida por (arriba por) la gráfica de $z = \varphi_2(x, y)$, y por (abajo por) la gráfica de $z = \varphi_1(x, y)$, y, posiblemente, por un "cilindro" que une a estas dos superficies sobre la región $R_1 \subset \mathbb{R}^2$.

Llamemos K a la frontera de Ω y supongamos que $\mathbf{f}$: $S \subset \mathbb{R}^2 \to \mathbb{R}^3$ es una parametrización de K, la cual produce una orientación con las normales a K apuntando *hacia* su *exterior*. De hecho, poniendo $K = K_1 \cup K_2 \cup K_3$, donde K_1 es la gráfica de $z = \varphi_1(x, y)$, K_2 es la gráfica de $z = \varphi_2(x, y)$ y K_3 es el (posible) cilindro que hay entre K_1 y K_2, tenemos que:

a. Una parametrización de K_1 (siendo ésta la gráfica de $z = \varphi_1(x, y)$ en la región R_1) es $\mathbf{f}$: $R_1 \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, \varphi_1(u, v))$. La orientación que ésta produce en K_1 es la dada por los vectores $\mathbf{N_f}(u, v)$ que son

$$\mathbf{N_f}(u, v) = \left(-\frac{\partial \varphi_1}{\partial u}, -\frac{\partial \varphi_1}{\partial v}, 1\right)$$

Nótese que estos vectores apuntan hacia el *interior* de K (¿por qué?: vea su tercera coordenada e imagine el dibujo correspondiente). Es decir, esta parametrización produce la orientación contraria a la deseada de K.

b. Una parametrización de K_2 (siendo ésta la gráfica de $z = \varphi_2(x, y)$ en la región R_1) es $\mathbf{f}$: $R_1 \subset \mathbb{R}^2 \to \mathbb{R}^3$,$\mathbf{f}(u, v) = (u, v, \varphi_2(u, v))$. La orientación que ésta produce en K_2 es la dada por los vectores $\mathbf{N_f}(u, v)$ que son

$$\mathbf{N_f}(u, v) = \left(-\frac{\partial \varphi_2}{\partial u}, -\frac{\partial \varphi_2}{\partial v}, 1\right)$$

los cuales apuntan hacia el exterior de K, como tiene que ocurrir.

Estas observaciones se aplicarán en la prueba del teorema de la divergencia que presentamos a continuación.

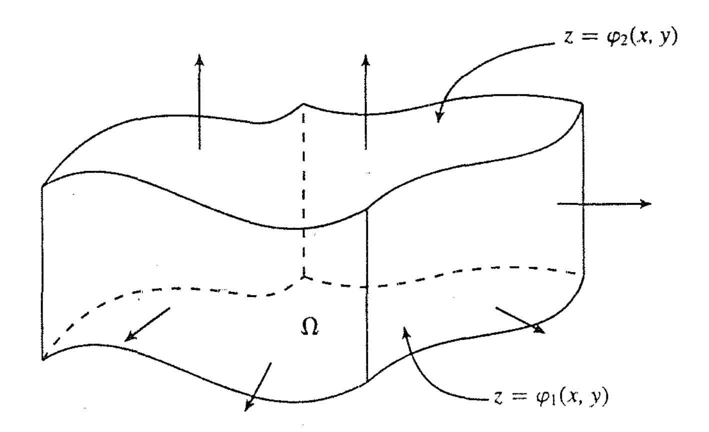


Figura 1. La región Ω y los vectores normales a K.

Teorema 9.3.1 (Teorema de la divergencia en el espacio). Sea Ω una región en $\mathbb{R}^3$ (del tipo mencionado en la exposición anterior) y sea K la frontera de Ω orientada con sus vectores normales apuntando hacia su exterior (digamos que parametrizada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$). Si $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$, es un campo vectorial de clase $\mathscr{C}^1$ definido en el abierto U que contiene a Ω, entonces

$$\iint_{K_\mathbf{f}} \mathbf{F}\, dA = \iiint_\Omega \operatorname{div} \mathbf{F} dx\, dy\, dz$$

Demostración. Por una parte se tiene

$$\begin{aligned}\iint_{K_\mathbf{f}} \mathbf{F}\, dA &= \iint_R \mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N}_\mathbf{f}(u, v) du\, dv \\ &= \iint_R F_1(\mathbf{f}(u, v)) \frac{\partial(y, z)}{\partial(u, v)} du\, dv + \iint_R F_2(\mathbf{f}(u, v)) \frac{\partial(z, x)}{\partial(u, v)} du\, dv \\ &+ \iint_R F_3(\mathbf{f}(u, v)) \frac{\partial(x, y)}{\partial(u, v)} du\, dv\end{aligned}$$

donde $\mathbf{f}(u, v) = (x(u, v), y(u, v), z(u, v))$. Por otra parte

$$\begin{aligned}\iiint_\Omega \operatorname{div} \mathbf{F} dx\, dy\, dz &= \iiint_\Omega \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \right) dx\, dy\, dz \\ &= \iiint_\Omega \frac{\partial F_1}{\partial x} dx\, dy\, dz + \iiint_\Omega \frac{\partial F_2}{\partial y} dx\, dy\, dz + \iiint_\Omega \frac{\partial F_3}{\partial z} dx\, dy\, dz\end{aligned}$$

La demostración del teorema queda completa si se prueban las 3 igualdades

$$\iint_R F_1(\mathbf{f}(u,v))\frac{\partial(y,z)}{\partial(u,v)}du\,dv = \iiint_\Omega \frac{\partial F_1}{\partial x}\,dx\,dy\,dz$$

$$\iint_R F_2(\mathbf{f}(u,v))\frac{\partial(z,x)}{\partial(u,v)}du\,dv = \iiint_\Omega \frac{\partial F_2}{\partial y}\,dx\,dy\,dz$$

$$\iint_R F_3(\mathbf{f}(u,v))\frac{\partial(x,y)}{\partial(u,v)}du\,dv = \iiint_\Omega \frac{\partial F_3}{\partial z}\,dx\,dy\,dz$$

El argumento que prueba cada una de estas igualdades es similar. Solamente presentaremos el que prueba la tercera de ellas. Como Ω es del tipo

$$\Omega = \{(x,y,z)|(x,y)\in R,\ \varphi_1(x,y)\le z\le \varphi_2(x,y)\}$$

podemos escribir

$$\begin{aligned}\iiint_\Omega \frac{\partial F_3}{\partial z}\,dx\,dy\,dz &= \iint_R\left(\int_{\varphi_1(x,y)}^{\varphi_2(x,y)}\frac{\partial F_3}{\partial z}dz\right)dxdy\\ &= \iint_R\big(F_3(x,y,\varphi_2(x,y)) - F_3(x,y,\varphi_1(x,y))\big)dxdy\end{aligned}$$

Por otra parte, poniendo $K = K_1\cup K_2\cup K_3$ como en la discusión previa al teorema, tenemos

$$\begin{aligned}\iint_R F_3(\mathbf{f}(u,v))\frac{\partial(x,y)}{\partial(u,v)}du\,dv &= \iint_K (0,0,F_3)\cdot\mathbf{N}\,dA\\ &= \iint_{K_1}(0,0,F_3)\cdot\mathbf{N}\,dA + \iint_{K_2}(0,0,F_3)\cdot\mathbf{N}\,dA + \iint_{K_3}(0,0,F_3)\cdot\mathbf{N}\,dA\end{aligned}$$

En el tercer sumando de la expresión anterior aparece el producto punto del campo $\tilde{F} = (0,0,F_3)$ con el vector $\mathbf{N}$, el cual apunta hacia el exterior de K. Recuerde que K_3 es la parte de la superficie K que conecta las gráficas de $z = \varphi_1(x,y)$ y $z = \varphi_2(x,y)$. Entonces ésta es un "cilindro", cuyos vectores normales son ***perpendiculares al eje*** z; es decir, $\mathbf{N}$ debe ser del tipo

$$\mathbf{N} = (a,b,0)$$

de modo que $\tilde{F}\cdot\mathbf{N} = 0$. Así pues, el tercer sumando es igual a cero. Trabajemos con el primer sumando: K_1 es una parte de K parametrizada por $\mathbf{f}_1\colon R\subset\mathbb{R}^2\to\mathbb{R}^3$, $\mathbf{f}_1(u,v) = (u,v,\varphi_1(u,v))$, la cual produce una orientación contraria a la establecida en el teorema. Se calcula la integral de

superficie correspondiente con esta parametrización, anteponiendo un signo menos. Es decir

$$\begin{aligned}\iint_{K_1} (0, 0, F_3) \cdot \mathbf{N}\, dA &= -\iint_{K_{1_{f_1}}} (0, 0, F_3) \cdot \mathbf{N}\, dA \\ &= -\iint_R (0, 0, F_3(\mathbf{f}_1(u, v))) \cdot \mathbf{N}_{\mathbf{f}_1}(u, v) du\, dv \\ &= -\iint_R F_3(\mathbf{f}_1(u, v)) \frac{\partial(u, v)}{\partial(u, v)} du\, dv = -\iint_R F_3(u, v, \varphi_1(u, v)) du\, dv\end{aligned}$$

Del mismo modo (para K_2), la función $\mathbf{f}_2 \colon R \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}_2(u, v) = (u, v, \varphi_2(u, v))$ parametriza la parte de la superficie K correspondiente a K_2 y además respeta la orientación establecida (hacia el exterior de K). Entonces

$$\begin{aligned}\iint_{K_2} (0, 0, F_3) \cdot \mathbf{N}\, dA &= \iint_{K_{2_{f_2}}} (0, 0, F_3) \cdot \mathbf{N}\, dA \\ &= \iint_R (0, 0, F_3(\mathbf{f}_2(u, v))) \cdot \mathbf{N}_{\mathbf{f}_2}(u, v) du\, dv \\ &= \iint_R F_3(\mathbf{f}_2(u, v)) \frac{\partial(u, v)}{\partial(u, v)} du\, dv = \iint_R F_3(u, v, \varphi_2(u, v)) du\, dv\end{aligned}$$

Entonces

$$\begin{aligned}\iint_R F_3(\mathbf{f}(u, v)) \frac{\partial(x, y)}{\partial(u, v)} du\, dv &= \iint_R F_3(u, v, \varphi_2(u, v)) du\, dv - \iint_R F_3(u, v, \varphi_1(u, v)) du\, dv \\ &= \iint_R \Big(F_3(u, v, \varphi_2(u, v)) - F_3(u, v, \varphi_1(u, v)) \Big) du\, dv \\ &= \iiint_\Omega \frac{\partial F_3}{\partial z} dx\, dy\, dz\end{aligned}$$

como se quería probar. Q.E.D.

Veamos algunos ejemplos.

Ejemplo 1. Un campo constante $\mathbf{F} \colon \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (\alpha, \beta, \gamma)$, tiene divergencia igual a cero (¿por qué?). Si K es una superficie cerrada como en las hipótesis del Teorema de la Divergencia, se tiene que el flujo de $\mathbf{F}$ a través de K es, según este teorema, igual a

$$\iint_{K_f} \mathbf{F}\, dA = \iiint_\Omega \operatorname{div} \mathbf{F}\, dx\, dy\, dz = \iiint_\Omega 0\, dx\, dy\, dz = 0$$

Esta fue la conclusión obtenida en los ejemplos 6 y 7 de la sección anterior, donde se calculó directamente el flujo de un campo constante a través de un elipsoide y de un cubo, respectivamente. ■

Ejemplo 2. Consideremos el campo radial $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x, y, z)$. La divergencia de este campo es

$$\operatorname{div} \mathbf{F} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 1 + 1 + 1 = 3$$

El flujo de este campo a través de la esfera $x^2 + y^2 + z^2 = c^2$, con sus normales apuntando hacia el exterior de ella, es, según el teorema de la divergencia, igual a

$$\iint_{K_f} \mathbf{F}\, dA = \iiint_{\Omega} \operatorname{div} \mathbf{F}\, dx\, dy\, dz = \iiint_{\Omega} 3\, dx\, dy\, dz$$

$$= 3 \iiint_{\Omega} dx\, dy\, dz = 3(\text{Volumen de la esfera}) = 3\left(\frac{4}{3}\pi c^3\right) = 4\pi c^3$$

que es el resultado obtenido en el ejemplo 3 de la sección anterior. ■

Ejemplo 3. Para el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x^3, y^3, z)$, tenemos

$$\operatorname{div} \mathbf{F} = \frac{\partial}{\partial x}(x^3) + \frac{\partial}{\partial y}(y^3) + \frac{\partial}{\partial z}(z) = 3x^2 + 3y^2 + 1$$

Calculemos el flujo de $\mathbf{F}$ a través de la esfera unitaria K, $x^2 + y^2 + z^2 = 1$ con sus normales apuntando hacia el exterior. Se tiene, según el teorema de la divergencia, que

$$\iint_{K_f} \mathbf{F}\, dA = \iiint_{\Omega} \operatorname{div} \mathbf{F}\, dx\, dy\, dz = \iiint_{\Omega} (3x^2 + 3y^2 + 1)\, dx\, dy\, dz$$

$$\underset{\longrightarrow}{\overset{\text{esfericas}}{=}} \int_0^{\pi} \int_0^{2\pi} \int_0^1 (3r^2 \cos^2\theta \operatorname{sen}^2\phi + 3r^2 \operatorname{sen}^2\theta \operatorname{sen}^2\phi + 1) r^2 \operatorname{sen}\phi\, dr\, d\theta\, d\phi$$

$$= 3 \int_0^{\pi} \operatorname{sen}^3\phi\, d\phi \int_0^{2\pi} d\theta \int_0^1 r^4 dr + \int_0^{\pi} \operatorname{sen}\phi\, d\phi \int_0^{2\pi} d\theta \int_0^1 r^2\, dr$$

$$= \frac{8}{5}\pi + \frac{4}{3}\pi = \frac{44}{15}\pi$$ ■

Con el teorema de la divergencia podemos dar un argumento sencillo que muestra que en efecto la divergencia de un campo (de clase $\mathscr{C}^1$) $\mathbf{F}$ en $\mathbb{R}^3$ se puede ver como el límite del cociente del flujo del campo a través del cuerpo (una superficie) K (no necesariamente un cubo) que circunda a $\mathbf{p}$, entre el volumen contenido por K, cuando K se contrae a $\mathbf{p}$. En efecto, sea K un cuerpo (digamos una superficie parametrizada por la función $\mathbf{f}$) que encierra una región Ω de $\mathbb{R}^3$ en la que se encuentra $\mathbf{p}$. El flujo de $\mathbf{F}$ a través de K es, según el teorema de la divergencia

$$\iint_{K_f} \mathbf{F}\, dA = \iiint_{\Omega} \operatorname{div} \mathbf{F}\, dx\, dy\, dz$$

Aplicando el teorema del valor medio a la integral triple de la derecha (sección 8.3 capítulo 6) se tiene

$$\iiint_{\Omega} \operatorname{div} \mathbf{F}\, dx\, dy\, dz = \operatorname{div} \mathbf{F}(\tilde{\mathbf{p}}) \iiint_{\Omega} dx\, dy\, dz = \operatorname{div} \mathbf{F}(\tilde{\mathbf{p}}) \text{ Volumen de } \Omega$$

donde $\tilde{\mathbf{p}}$ es un punto de Ω. Entonces

$$\operatorname{div} \mathbf{F}(\tilde{\mathbf{p}}) = \frac{\iint_{K_f} \mathbf{F}\, dA}{\text{Volumen de } \Omega} = \frac{\text{Flujo de } \mathbf{F} \text{ a través de } K}{\text{Volumen encerrado por } K}$$

Haciendo que la superficie K se contraiga hacia el punto $\mathbf{p}$, y usando la continuidad de las derivadas parciales que aparecen en la expresión de la divergencia del campo, nos queda

$$\lim_{K \to \mathbf{p}} \operatorname{div} \mathbf{F}(\tilde{\mathbf{p}}) = \operatorname{div} \mathbf{F}(\mathbf{p}) = \lim_{K \to \mathbf{p}} \frac{\text{Flujo de } \mathbf{F} \text{ a través de } K}{\text{Volumen encerrado por } K}$$

como queríamos establecer.

El teorema de la divergencia es válido también para superficies K que contienen regiones más generales en $\mathbb{R}^3$, por ejemplo, las que se pueden dividir en varias subregiones como las consideradas por el teorema. Aunque no entraremos en detalles al respecto, simplemente comentamos cómo es que esta generalización puede parecer natural.

Supongamos que la superficie cerrada K orientada con sus vectores normales apuntando hacia su exterior, o sea parametrizada por la función $\mathbf{f}$, contiene una región Ω que puede ser presentada como la unión de dos subregiones Ω_1 y Ω_2 que son de los tres tipos mencionados en la discusión previa al teorema 9.3.1. Sean K_1 y K_2 las superficies correspondientes a las fronteras de Ω_1 y Ω_2, orientadas con sus vectores normales apuntando hacia su exterior (digamos que parametrizadas por las funciones $\mathbf{f}_1$ y $\mathbf{f}_2$, respectivamente). Es claro que la "frontera común" de Ω_1 y Ω_2 va a ser parte tanto de K_1 como de K_2, con la diferencia de que los vectores normales en cada una de éstas estarán en direcciones contrarias.

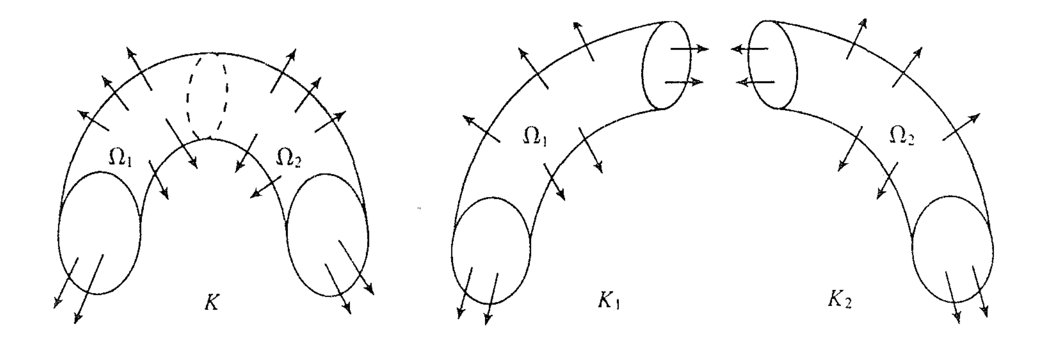


Figura 2. La región Ω divida en dos subregiones y Ω_1 y Ω_2.

Con este teorema podemos calcular el flujo de $\mathbf{F}$ a través de K_1 y K_2, puesto que estas superficies contienen regiones en $\mathbb{R}^3$ que cumplen con las hipótesis del teorema de la divergencia.

$$\text{Flujo de } \mathbf{F} \text{ a través de } K_1 = \iint_{K_{1_{f_1}}} \mathbf{F}\, dA = \iiint_{\Omega_1} \operatorname{div} \mathbf{F}\, dx\, dy\, dz$$

$$\text{Flujo de } \mathbf{F} \text{ a través de } K_2 = \iint_{K_{2_{f_2}}} \mathbf{F}\, dA = \iiint_{\Omega_2} \operatorname{div} \mathbf{F}\, dx\, dy\, dz$$

Nótese ahora que el flujo de **F** a través de K se puede calcular como la suma del flujo de **F** a través de K_1 más el flujo a través de K_2, pues en la frontera común los flujos correspondientes se anularán uno al otro. Así pues

$$\iint_{K_f} \mathbf{F}\, dA = \text{Flujo de } \mathbf{F} \text{ a través de} K$$

$$= \text{Flujo de } \mathbf{F} \text{ a través de} K_1 + \text{Flujo de } \mathbf{F} \text{ a través de} K_2$$

$$= \iint_{K_{1_{f_1}}} \mathbf{F}\, dA + \iint_{K_{2_{f_2}}} \mathbf{F}\, dA = \iiint_{\Omega_1} \operatorname{div} \mathbf{F} dx\, dy + \iiint_{\Omega_2} \operatorname{div} \mathbf{F}\, dx\, dy$$

$$= \iiint_{\Omega} \operatorname{div} \mathbf{F} dx\, dy$$

Es decir, vale la fórmula del teorema de la divergencia

$$\iint_{K_f} \mathbf{F}\; dA = \iiint_{\Omega} \operatorname{div} \mathbf{F}\, dx\, dy$$

para la superficie K que contiene la región Ω.

El siguiente importante resultado usará para su demostración el teorema de la divergencia en esta versión con superficies más generales.

Teorema (Teorema de Gauss). Sea Ω una región en $\mathbb{R}^3$ (del tipo mencionado en la discusión previa la teorema 9.3.1) y sea K la frontera de Ω orientada con sus normales apuntando hacia su exterior (digamos que parametrizada por la función $\mathbf{f}\colon R \subset \mathbb{R}^2 \to \mathbb{R}^3$). Suponga que $(0, 0, 0) \notin K$. Sea $\mathbf{F}\colon \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}^3$ el campo vectorial

$$\mathbf{F}(x, y, z) = \frac{1}{\|(x, y, z)\|^3}(x, y, z)$$

Entonces el flujo de **F** a través de K es

$$\iint_{K_f} \mathbf{F}\, dA = \begin{cases} 4\pi_0 & \text{si } (0, 0, 0) \in \Omega \\ 0 & \text{si } (0, 0, 0) \notin \Omega \end{cases}$$

Demostración. Nótese que el campo dado tiene por divergencia

$$\begin{aligned} \operatorname{div} \mathbf{F} &= \frac{\partial}{\partial x}\left(\frac{x}{(x^2 + y^2 + z^2)^{3/2}}\right) + \frac{\partial}{\partial y}\left(\frac{y}{(x^2 + y^2 + z^2)^{3/2}}\right) + \frac{\partial}{\partial z}\left(\frac{z}{(x^2 + y^2 + z^2)^{3/2}}\right) \\ &= \frac{y^2 + z^2 - 2x^2}{(x^2 + y^2 + z^2)^{5/2}} + \frac{x^2 + z^2 - 2y^2}{(x^2 + y^2 + z^2)^{5/2}} + \frac{x^2 + y^2 - 2z^2}{(x^2 + y^2 + z^2)^{5/2}} = 0 \end{aligned}$$

Entonces, si $(0, 0, 0) \notin \Omega$, se tienen todas las hipótesis del teorema de la divergencia satisfechas (el dominio del campo **F**, $\mathbb{R}^3 - \{(0, 0, 0)\}$, es un abierto de $\mathbb{R}^3$ que contiene a K), de modo que en este

caso, con tal teorema, se tiene

$$\iint_{K_f} F\,dA = \iiint_{\Omega} \operatorname{div} F \; dx\,dy\,dz = \iiint_{\Omega} 0 \; dx\,dy\,dz = 0$$

Consideremos entonces el caso en que $(0, 0, 0) \in \Omega$. En este caso ya no es cierto que el dominio de $\mathbf{F}$ ($\mathbb{R}^3 - \{(0, 0, 0)\}$) *contenga* a K, ya que en la región que K encierra se encuentra el punto $(0, 0, 0)$. Sea B_ϵ una bola con centro en $(0, 0, 0)$ y radio $\epsilon > 0$ que se encuentra contenida en Ω (de modo más preciso, en el interior de Ω, el cual es un conjunto abierto, lo que garantiza, a su vez, que dicha $\epsilon > 0$ exista). Es decir

$$B_\epsilon = \{(x, y, z) | x^2 + y^2 + z^2 \leq \epsilon^2\} \subset \Omega$$

Consideremos la región $\Omega^* = \Omega \backslash B_\epsilon$. Esta es entonces la región Ω a la que le hemos quitado una pequeña bola que contiene el origen de coordenadas. Sea K^* la frontera de Ω^* con sus normales apuntando hacia el exterior. Obsérvese que K^* está constituido por K (la frontera de Ω) unión la frontera de la bola B_ϵ. Las normales en K (como parte de K^*) apuntan hacia el exterior de K (que es también el exterior de K^*), y las normales en ∂B_ϵ, frontera de B_ϵ, deberán apuntar hacia el *interior* de B_ϵ (es decir, hacia el origen), el cual es también parte del exterior de K^*.

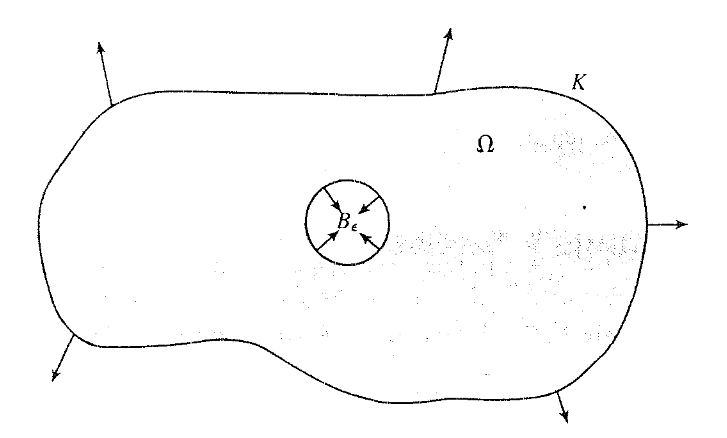


Figura 3. La región Ω^*, la bola B_ϵ, la superficie K y las normales a K.

Un campo de vectores normales a ∂B_ϵ que apunta hacia su interior es $\mathbf{N}\colon \partial B_\epsilon \to \mathbb{R}^3$,

$$\mathbf{N}(x, y, z) = \frac{-1}{\|(x, y, z)\|}(x, y, z) = -\frac{1}{\epsilon}(x, y, z)$$

Entonces el flujo de $\mathbf{F}$ a través de ∂B_ϵ es

$$\iint_{\partial B_\epsilon} F \cdot \mathbf{N}\,dA = \iint_{\partial B_\epsilon} \frac{1}{\|(x, y, z)\|^3}(x, y, z) \cdot \left(\frac{-1}{\|(x, y, z)\|}(x, y, z)\right) dA$$

$$= -\iint_{\partial B_\epsilon} \frac{1}{\|(x, y, z)\|^4} \|(x, y, z)\|^2\,dA = -\iint_{\partial B_\epsilon} \frac{1}{\|(x, y, z)\|^2}\,dA$$

En los puntos ∂B_ϵ se tiene $\|(x, y, z)\| = \epsilon$, de modo que

$$-\iint_{\partial B_\epsilon} \frac{1}{\|(x, y, z)\|^2}\, dA = -\iint_{\partial B_\epsilon} \frac{1}{\epsilon^2}\, dA = -\frac{1}{\epsilon^2} \iint_{\partial B_\epsilon} dA$$

$$= -\frac{1}{\epsilon^2}(\text{area de } \partial B_\epsilon) = -\frac{1}{\epsilon^2}(4\pi\epsilon^2) = -4\pi$$

Aplicando el teorema de la divergencia con la superficie K^* (parametrizada digamos por la función f^*), el cual ahora sí se puede usar (en su versión generalizada, según la discusión previa a este teorema) pues el dominio $\mathbb{R}^3 - \{(0, 0, 0)\}$ contiene a K^*, tenemos

$$\iint_{K^*_{f^*}} \mathbf{F}\, dA = \iiint_{\Omega^*} \operatorname{div} \mathbf{F}\, dx\, dy\, dz = \iiint_{\Omega^*} 0\, dx\, dy\, dz = 0$$

$$= \iint_K \mathbf{F} \cdot \mathbf{N}\, dA + \iint_{\partial B_\epsilon} \mathbf{F} \cdot \mathbf{N}\, dA = \iint_K \mathbf{F} \cdot \mathbf{N}\, dA - 4\pi$$

Entonces el flujo de $\mathbf{F}$ a través de K es

$$\iint_{K_f} \mathbf{F}\, dA = \iint_K \mathbf{F} \cdot \mathbf{N}\, dA = 4\pi$$

como se quería demostrar. Q.E.D.

Ejercicios (Capítulo 9, Sección 3)

En los ejercicios 1—5, verifique el teorema de la divergencia con el campo $\mathbf{F}$ dado y la superficie K indicada.

1. $\mathbf{F}(x, y, z) = (y, x, z)$, $K = \{(x, y, z)|x^2 + y^2 + z^2 = 1\}$.

2. $\mathbf{F}(x, y, z) = (xy, xz, yz)$, K es el tetraedro formado por los planos coordenados y la porción del plano $x + y + z = 1$ que se encuentra en el primer octante.

3. $\mathbf{F}(x, y, z) = (x + y, y + z, z + x)$, K es el cubo $[0, 1]^3$.

4. $\mathbf{F}(x, y, z) = (xyz, 1, 1)$, K es el elipsoide $x^2 + 2y^2 + 3z^2 = 6$.

5. $\mathbf{F}(x, y, z) = (z^2, y^2, x^2)$, K es la octava parte de la esfera $x^2 + y^2 + z^2 = 1$ que se encuentra en el primer octante, junto con los planos coordenados.

6. Con el teorema de la divergencia calcule el flujo del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x^3, y^3, z^3)$ a través de la esfera unitaria $x^2 + y^2 + z^2 = 1$.

7. Sea $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$ un campo de clase $\mathscr{C}^1$ tal que $(x, y, z) \cdot \mathbf{F}(x, y, z) = 0$, es decir, $xF_1(x, y, z) + yF_2(x, y, z) + zF_3(x, y, z) = 0$ para todo $(x, y, z) \in \mathbb{R}^3$.

a. Describa de modo geométrico el campo $\mathbf{F}$.

b. Demuestre que el flujo de $\mathbf{F}$ a través de la esfera $x^2 + y^2 + z^2 = c^2$ es igual a cero.

8. Suponga que el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$ tiene divergencia igual a 0. Demuestre que el flujo de $\mathbf{F}$ a través de cualquier superficie cerrada (de las consideradas en el teorema 9.3.1) es igual a cero.

9. Suponga que el flujo de un campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$ a través de una superficie cerrada K es igual a cero. ¿Se concluye de aquí que la divergencia del campo es cero?

10. El campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ está descrito, en coordenadas cilíndricas, como $\mathbf{F}(r, \theta, z) = r^2\mathbf{e}_r + r\mathbf{e}_\theta + z^2\mathbf{e}_z$. Use el teorema de la divergencia para calcular el flujo de este campo a través:
   a. de la esfera $x^2 + y^2 + z^2 = 1$,
   b. del cilindro $K = \{(x, y, z)|x^2 + y^2 = 1, -1 \leq z \leq 1\} \cup \{(x, y, z)|x^2 + y^2 \leq 1, z = \pm 1\}$ (Ver apéndice de la sección 11 del capítulo 7).

11. Repita el ejercicio anterior con el campo $\mathbf{F}(r, \theta, z) = r\cos\theta\mathbf{e}_r + r\operatorname{sen}\theta\mathbf{e}_\theta + z\mathbf{e}_z$.

12. El campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ viene dado, en coordenadas esféricas, como $\mathbf{F}(r, \theta, \phi) = r\operatorname{sen}\phi\mathbf{e}_r + \cos\phi\mathbf{e}_\theta + r^2\operatorname{sen}\phi\mathbf{e}_\phi$. Con el teorema de la divergencia calcule el flujo del campo $\mathbf{F}$ a través:
   a. de la esfera $x^2 + y^2 + z^2 = 1$,
   b. del cono $K = \{(x, y, z)|z^2 = x^2 + y^2, z \geq 0\} \cup \{(x, y, z)|x^2 + y^2 \leq 1, z = 1\}$

13. Repita el ejercicio anterior con el campo $\mathbf{F}(r, \theta, \phi) = r\mathbf{e}_r + \phi\mathbf{e}_\theta$.

14. El campo eléctrico producido por una carga positiva $\mathbf{q}$ que se encuentra en el origen de coordenadas, es $\mathbf{E}: \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}^3$

$$\mathbf{E}(x, y, z) = \frac{k\mathbf{q}}{\|(x, y, z)\|^3}(x, y, z)$$

donde k es la constante de proporcionalidad de la ley de Coulomb (ver sección 8 del capítulo 7). Con un sistema de unidades adecuado podemos suponer que $k = 1$. Demuestre que el flujo de este campo a través de una superficie cerrada (de las consideradas en el teorema 9.3.1) es igual a 4π veces el valor de la carga $\mathbf{q}$, si ésta se encuentra dentro de la superficie, y cero, en caso de que la carga se encuentre fuera.

9.4 El rotacional de un campo vectorial

En esta sección introducimos otro concepto importante en el estudio de campos vectoriales, llamado rotacional (que, junto con los conceptos de gradiente y divergencia, constituyen la base del "lenguaje de operación" en el estudio de campos vectoriales; una visión conjunta de estas operaciones se dará en la sección 6).

Supongamos un campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_x, F_y, F_z)$, diferenciable, definido en el conjunto abierto U de $\mathbb{R}^3$. Se define el *rotacional* de $\mathbf{F}$ en el punto $\mathbf{p}$ de U, denotado por rot $\mathbf{F}(\mathbf{p})$, como

$$\operatorname{rot}\mathbf{F}(\mathbf{p}) = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)$$

en que las derivadas parciales de las funciones componentes de $\mathbf{F}$, $F_x, F_y, F_z: U \subseteq \mathbb{R}^3 \to \mathbb{R}$, se evalúan en el punto $\mathbf{p}$. Obsérvese que rot $\mathbf{F}(\mathbf{p})$ es un vector en $\mathbb{R}^3$, de modo que podemos hablar del

"campo rotacional" de $\mathbf{F}$, rot $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, el cual a cada punto $\mathbf{p} \in U$ le asocia el vector rot $\mathbf{F}(\mathbf{p}) \in \mathbb{R}^3$. Una manera sencilla de recordar la fórmula que define este campo, es por medio del determinante

$$\det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$

del cual, si hacemos un desarrollo "formal" respecto de su primera línea, entendiendo que la "multiplicación del símbolo $\frac{\partial}{\partial x}$ por el símbolo F_y es igual a la derivada parcial $\frac{\partial F_y}{\partial x}$ " (de modo semejante para las otras combinaciones de estos símbolos), se obtiene

$$\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\mathbf{k}$$

que es el vector que escribimos anteriormente para definir rot $\mathbf{F}(\mathbf{p})$. Así pues, podemos escribir de manera formal

$$\text{rot } F = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$ ■

Ejemplo 1. El rotacional del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x^2y, x + 3y + z, xyz^3)$ es

$$\begin{aligned}
\text{rot } \mathbf{F} &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ x^2y & x + 3y + z & xyz^3 \end{bmatrix} \\
&= \left(\frac{\partial}{\partial y}(xyz^3) - \frac{\partial}{\partial z}(x + 3y + z)\right)\mathbf{i} + \left(\frac{\partial}{\partial z}(x^2y) - \frac{\partial}{\partial x}(xyz^3)\right)\mathbf{j} \\
&+ \left(\frac{\partial}{\partial x}(x + 3y + z) - \frac{\partial}{\partial y}(x^2y)\right)\mathbf{k} \\
&= (xz^3 - 1)\mathbf{i} + (0 - yz^3)\mathbf{j} + (1 - x^2)\mathbf{k} \\
&= (xz^3 - 1, -yz^3, 1 - x^2)
\end{aligned}$$ ■

Ejemplo 2. El rotacional del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (\alpha xy, \beta x, 0)$, es

$$\text{rot } \mathbf{F} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \alpha xy & \beta x & 0 \end{bmatrix} = (0, 0, \beta - \alpha x)$$ ■

El rotacional de un campo $\mathbf{F}$ en $\mathbb{R}^3$ es un concepto que generaliza al de rotación de un campo en $\mathbb{R}^2$, estudiado en la sección 10 del capítulo 7. Recuerde que la rotación de un campo continuo se definió $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ alrededor de $\boldsymbol{\lambda}$ (un camino cerrado simple de clase $\mathscr{C}^1$), denotada por rtc $\mathbf{F}(\boldsymbol{\lambda})$, como

$$\text{rtc } \mathbf{F}(\boldsymbol{\lambda}) = \frac{1}{\text{Área contenida por } \boldsymbol{\lambda}} \int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$$

donde a $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ se le había llamado circulación del campo $\mathbf{F}$ alrededor de $\boldsymbol{\lambda}$. Así pues, rtc $\mathbf{F}(\boldsymbol{\lambda})$ es una medida de la circulación del campo alrededor de $\boldsymbol{\lambda}$, por unidad de área. Para hacer este concepto "puntual", se tomaban todos los caminos circulares $\boldsymbol{\lambda}_\epsilon$ de radio $\epsilon > 0$ con centro en el punto $\mathbf{p} \in U$, y se calculaba la rotación de $\mathbf{F}$ a través de $\mathbf{F}$ alrededor de cada uno de estos caminos, definiéndose la rotación de $\mathbf{F}$ en el punto $\mathbf{p} \in U$ como el límite de la rotación de $\mathbf{F}$ alrededor de $\boldsymbol{\lambda}_\epsilon$ cuando ϵ tendía a cero. Es decir

$$\operatorname{rtc} \mathbf{F}(\mathbf{p}) = \lim_{\epsilon \to 0} \operatorname{rtc} \mathbf{F}(\boldsymbol{\lambda}_\epsilon)$$

Con la ayuda del teorema de Green se pudo establecer que si el campo $\mathbf{F}$ es de clase $\mathscr{C}^1$, su rotación en $\mathbf{p}$ se calculaba como

$$\operatorname{rtc} \mathbf{F}(\mathbf{p}) = \frac{\partial F_2}{\partial x}(\mathbf{p}) - \frac{\partial F_1}{\partial y}(\mathbf{p})$$

donde $\mathbf{F} = (F_1, F_2)$.

Se vio que rtc $\mathbf{F}(\mathbf{p})$ es un escalar que nos da información sobre la capacidad que tiene el campo $\mathbf{F}$ (viendo éste como campo de velocidades de un fluido en movimiento) de hacer girar un cuerpo (cuya forma estaba dada por la traza del camino $\boldsymbol{\lambda}$) que se encuentra en el flujo descrito por $\mathbf{F}$. Todas estas ideas se pueden adecuar para ver que el rotacional de un campo en $\mathbb{R}^3$ tiene también información sobre la circulación del campo por unidad de área en cada punto. Sin embargo, no lo haremos en este momento: será cuando tengamos establecido el teorema de Stokes (sección 6) que veremos, con la ayuda de éste, que el rotacional de $\mathbf{F}$ es, en cierto sentido que será precisado en su momento, una generalización a campos en $\mathbb{R}^3$ del concepto de rotación de campos en $\mathbb{R}^2$, que rescata todo el contenido físico de éste último.

Nótese, sin embargo, que mientras la rotación de un campo $\tilde{\mathbf{F}}$ en $\mathbb{R}^2$ es *un escalar*, el rotacional de un campo $\mathbf{F}$ en $\mathbb{R}^3$ (en cada punto de su dominio) *es un vector en* $\mathbb{R}^3$. Para ver el parentesco entre el escalar rtc $\tilde{\mathbf{F}}(\mathbf{p})$ y el vector rot $\mathbf{F}(\mathbf{p})$, considere el campo $\mathbf{F}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, de clase $\mathscr{C}^1$, definido en el conjunto $U = \{(x, y, z) | (x, y) \in R, z = 0\}$, donde R es un conjunto abierto de $\mathbb{R}^2$,

$$\mathbf{F}(x, y, 0) = (F_x(x, y, 0), F_y(x, y, 0), 0)$$

Es claro que podemos identificar a este campo con el campo en $\mathbb{R}^2$, $\tilde{\mathbf{F}}\colon R \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\tilde{\mathbf{F}}(x, y) = (\tilde{F}_x(x, y), \tilde{F}_y(x, y))$, donde $\tilde{\mathbf{F}}(x, y) = \mathbf{F}(x, y, 0)$. El rotacional de $\mathbf{F}$ (en un punto $\mathbf{p} = (x, y, 0) \in U$) es

$$\operatorname{rot} \mathbf{F} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & 0 \end{bmatrix} = \left(0, 0, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)$$

Se tiene pues que rot $\mathbf{F}(\mathbf{p})$ es un vector en $\mathbb{R}^3$ perpendicular al plano xy (donde se encuentra el campo $\mathbf{F}$), cuya coordenada z es

$$\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = \frac{\partial \tilde{\mathbf{F}}_y}{\partial x} - \frac{\partial \tilde{\mathbf{F}}_x}{\partial y} = \operatorname{rtc} \tilde{\mathbf{F}}$$

Así pues, rot $\mathbf{F}$ resulta ser el vector $(0, 0, \operatorname{rtc} \tilde{\mathbf{F}})$. Vea, por ejemplo el campo del ejemplo 2 (su rotacional) y compare con los ejemplos 5 y 6 de la sección 10 del capítulo 7.

Otro punto interesante sobre el rotacional de un campo $\mathbf{F}$ en $\mathbb{R}^3$ es la relación de éste con la propiedad de que el campo sea conservativo. En la sección 4 del capítulo 7 (Teorema 7.4.2) se vio una condición necesaria para que el campo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$, de clase $\mathscr{C}^k$, $k \geq 1$, definido en el conjunto abierto U de $\mathbb{R}^n$, sea conservativo es que

$$\frac{\partial F_i}{\partial x_j}(\mathbf{p}) = \frac{\partial F_j}{\partial x_i}(\mathbf{p})$$

para $\mathbf{p} \in U$, $1 \leq i < j \leq n$. Para el campo en $\mathbb{R}^3$, $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_x, F_y, F_z)$, de clase $\mathscr{C}^1$, esta condición se ve como

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}, \quad \frac{\partial F_x}{\partial z} = \frac{\partial F_z}{\partial x}, \quad \frac{\partial F_y}{\partial z} = \frac{\partial F_z}{\partial y}$$

lo cual se puede escribir como

$$\operatorname{rot} \mathbf{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) = (0, 0, 0)$$

A un campo $\mathbf{F}$ en $\mathbb{R}^3$ cuyo rotacional siempre es cero, se le llama (campo) *irrotacional*. Así pues, el teorema 7.4.2 se puede enunciar, para campos en $\mathbb{R}^3$, diciendo que una condición necesaria para que el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ sea conservativo es que este campo sea irrotacional. Es decir, se tiene la implicación

$$\mathbf{F} \text{ conservativo} \Rightarrow \mathbf{F} \text{ irrotacional}$$

Sabemos que la implicación recíproca es falsa en general, pues, por ejemplo, el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$

$$\mathbf{F}(x, y, z) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0 \right)$$

definido en el conjunto abierto $U = \{(x, y, z) | (x, y) \neq (0, 0)\} \subseteq \mathbb{R}^3$, es irrotacional, pero no es conservativo. Para ver la validez de esta afirmación consideremos la siguiente situación más general: en un campo en $\mathbb{R}^2$, $\tilde{\mathbf{F}}: R \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\tilde{\mathbf{F}} = (\tilde{F}_x, \tilde{F}_y)$, de clase $\mathscr{C}^1$, definido en el conjunto abierto R de $\mathbb{R}^2$, asociamos a éste el campo en $\mathbb{R}^3$, $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ definido en el conjunto $U = R \times \mathbb{R} = \{(x, y, z) | (x, y) \in R\}$ (obsérvese que U es un conjunto abierto de $\mathbb{R}^3$),

$$\mathbf{F}(x, y, z) = (F_x(x, y, z), F_y(x, y, z), F_z(x, y, z)) = (\tilde{F}_x(x, y), \tilde{F}_y(x, y), 0)$$

Ciertamente $\mathbf{F}$ es de clase $\mathscr{C}^1$. De hecho hemos verificado ya que

$$\operatorname{rot} \mathbf{F}(x, y, z) = (0, 0, \operatorname{rtc} \tilde{\mathbf{F}}(x, y))$$

Veremos que el campo $\mathbf{F}$ es conservativo si y sólamente si el campo $\tilde{\mathbf{F}}$ lo es. En efecto, si $\tilde{\mathbf{F}}$ es conservativo hay una función $\tilde{f}: R \subseteq \mathbb{R}^2 \to \mathbb{R}$ (de clase $\mathscr{C}^2$) de modo que

$$\operatorname{grad} \tilde{f}(x, y) = (\tilde{F}_x(x, y), \tilde{F}_y(x, y))$$

Defina $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ como $f(x, y, z) = \tilde{f}(x, y)$. Es claro que grad $f(x, y, z) = \mathbf{F}(x, y, z)$; es decir, **f** es una función potencial de **F**, y así, **F** es conservativo. En forma recíproca, suponga que **F** es conservativo. Hay entonces una función potencial $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ (de clase $\mathscr{C}^2$) tal que

$$\operatorname{grad} f(x, y, z) = \mathbf{F}(x, y, z) = (\tilde{F}_x(x, y), \tilde{F}_y(x, y), 0)$$

Entonces $\frac{\partial f}{\partial z}(x, y, z) = 0\ \forall (x, y, z) \in U$. Esto significa que *$f$ no depende de z* (asegúrese de que esta afirmación es verdadera). Es decir, se da una función $\tilde{f}: R \subseteq \mathbb{R}^2 \to \mathbb{R}$ que $f(x, y, z) = \tilde{f}(x, y)$. Se verifica fácilmente que grad $\tilde{f}(x, y) = \tilde{\mathbf{F}}(x, y)$. Es decir, que $\tilde{f}$ es una función potencial de $\tilde{\mathbf{F}}$ y que, por lo tanto, $\tilde{\mathbf{F}}$ es conservativo.

En nuestro caso tenemos el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$

$$\mathbf{F}(x, y, z) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0\right)$$

definido en el conjunto abierto $U = \{(x, y, z) | (x, y) \neq (0, 0)\} = (\mathbb{R}^2 - \{(0, 0)\}) \times \mathbb{R}$. Este es el campo asociado a $\tilde{\mathbf{F}}: \mathbb{R}^2 - \{(0, 0)\} \to \mathbb{R}^2$, $\tilde{\mathbf{F}}(x, y)\left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$, el cual sabemos desde el capítulo 7, que no es conservativo. Entonces **F** tampoco lo es, como queríamos ver.

En general, el hecho de que un campo sea irrotacional sólo nos permite concluir que el campo es localmente conservativo, es decir, cada punto de su dominio es el centro de una bola en la que está definida una función potencial para el campo (o bien, una bola tal que restingiendo el campo a ella, éste es conservativo). Este es, de hecho, el contenido del teorema 7.4.3 en su versión para campos en $\mathbb{R}^3$ con la nueva nomenclatura establecida en esta sección.

La validez de la equivalencia entre las propiedades de un campo de ser conservativo y ser irrotacional, se tiene en el caso (y solamente en este caso) en el que el dominio $U \subseteq \mathbb{R}^3$ del campo, sea un conjunto *simplemente conexo* (propiedad que falla en el conjunto $U = \{(x, y, z) | (x, y) \neq (0, 0)\}$ del campo **F** considerado anteriormente: este conjunto es el espacio $\mathbb{R}^3$ al que se la ha retirado el eje z, y entonces U no es simplemente conexo —¿por qué?). Este es el resultado establecido en el teorema 7.4.5.

La implicación: **F** conservativo $\Rightarrow$ **F** irrotacional, para un campo **F** en $\mathbb{R}^3$, tiene todavía otro resultado general inmerso: supongamos que $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ es una función de clase $\mathscr{C}^2$ definida en el conjunto abierto U de $\mathbb{R}^3$. Entonces su gradiente grad $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, grad $f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$ es, obviamente, un campo (de clase $\mathscr{C}^1$) conservativo. Es entonces irrotacional. Esta conclusión, vista así, es una consecuencia inmediata del teorema de Schwarz sobre la igualdad de las derivadas parciales de segundo orden mixtas para una función de clase $\mathscr{C}^2$, pues

$$\begin{aligned}
\operatorname{rot} \operatorname{grad} f &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \end{bmatrix} \\
&= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y}, \frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z}, \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}\right) \\
&= (0, 0, 0)
\end{aligned}$$

Así pues, para toda función $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ de clase $\mathscr{C}^2$, se tiene

$$\operatorname{rot} \operatorname{grad} f = 0$$

Apéndice El rotacional en los sistemas de coordenadas cilíndricas y esféricas

En este apéndice se establecerán las expresiones correspondientes al rotacional de un campo, cuando éstos (el campo dado y su campo rotacional) se ven en los sistemas de coordenadas cilíndricas y esféricas. En el apéndice de la sección 11 del capítulo 7 se presentaron las expresiones correspondientes a la divergencia de un campo en los sistemas de coordenadas cilíndricas y esféricas. En esa ocasión se dedujo, a partir de la concepción física de lo que es la divergencia, la expresión correspondiente a ésta en coordenadas cilíndricas, además de que dicha expresión se verificó posteriormente en forma directa. El caso de las coordenadas esféricas se dejó al lector para su verificación. Lo que haremos en este apéndice con el rotacional de un campo será distinto: dejaremos el caso de las coordenadas cilíndricas para que el lector lo trabaje (éste es el caso "simple"), y abordaremos el problema de *deducir* (y no de verificar) la expresión del rotacional de un campo en coordenadas esféricas (éste caso es, técnicamente, mucho más complicado que el caso de las coordenadas cilíndricas).

De este modo, sólo mencionaremos la situación que se presenta en el caso de las coordenadas cilíndricas. Sea $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ un campo de clase $\mathscr{C}^1$ (cuya expresión en el sistema cartesiano se ve como) $\mathbf{F} = (F_x, F_y, F_z) = F_x\mathbf{i} + F_y\mathbf{y} + F_z\mathbf{k}$. Sea $\tilde{\mathbf{F}}(r, \theta, z)$ la expresión del campo $\mathbf{F}$ en coordenadas cilíndricas. Sabemos que (ver inciso *a*) de la proposición del apéndice de la sección 2 del capítulo 7)

$$\begin{aligned}\mathbf{F}(r\cos\theta, r\operatorname{sen}\theta, z) &= \tilde{\mathbf{F}}(r, \theta, z) = (F_r(r, \theta, z), F_\theta(r, \theta, z), F_z(r, \theta, z)) \\ &= F_r(r, \theta, z)\mathbf{e}_r + F_\theta(r, \theta, z)\mathbf{e}_\theta + F_z(r, \theta, z)\mathbf{e}_z\end{aligned}$$

donde

$$\begin{aligned}F_r(r, \theta, z) &= \cos\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z) + \operatorname{sen}\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z) \\ F_\theta(r, \theta, z) &= -\operatorname{sen}\theta F_x(r\cos\theta, r\operatorname{sen}\theta, z) + \cos\theta F_y(r\cos\theta, r\operatorname{sen}\theta, z) \\ F_z(r, \theta, z) &= F_z(r\cos\theta, r\operatorname{sen}\theta, z)\end{aligned}$$

La expresión del rotacional del campo $\tilde{\mathbf{F}}(r, \theta, z)$ es

$$\operatorname{rot}\tilde{\mathbf{F}} = \frac{1}{r}\left(\frac{\partial F_z}{\partial\theta} - \frac{\partial}{\partial z}(rF_\theta)\right)\mathbf{e}_r + \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right)\mathbf{e}_\theta + \frac{1}{r}\left(\frac{\partial}{\partial r}(rF_\theta) - \frac{\partial F_r}{\partial\theta}\right)\mathbf{e}_z$$

o bien, de manera formal (el desarrollo respecto de la primera línea del determinante)

$$\boxed{\operatorname{rot}\tilde{\mathbf{F}} = \frac{1}{r}\det\begin{bmatrix}\mathbf{e}_r & r\mathbf{e}_\theta & \mathbf{e}_z \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial z} \\ F_r & rF_\theta & F_z\end{bmatrix}}$$

donde $\operatorname{rot}\tilde{\mathbf{F}}$ y las derivadas parciales involucradas están evaluadas en un punto (r, θ, z) arbitrario. Como decíamos, esta expresión puede ser verificada directamente (con sólo hacer las "operaciones" que aparecen en $\operatorname{rot}\tilde{\mathbf{F}}$ y viendo que éstas producen la expresión que ya conocemos de $\operatorname{rot}\mathbf{F}$).

Consideremos entonces el problema de escribir el rotacional de un campo en coordenadas esféricas: tenemos un campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ de clase $\mathscr{C}^1$, $\mathbf{F} = (F_x, F_y, F_z) = F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k}$.

Se trata entonces de obtener la expresión del campo rot $\tilde{\mathbf{F}}$ en coordenadas esféricas, donde $\tilde{\mathbf{F}}$ es la expresión del campo $\mathbf{F}$ en este nuevo sistema coordenado. La expresión de $\tilde{\mathbf{F}}$ ya la conocemos (ver inciso b) de la proposición del apéndice de la sección 2 del capítulo 7). Se tiene

$$\begin{aligned}\mathbf{F}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi) = \tilde{\mathbf{F}}(r,\theta,\phi) &= (F_r(r,\theta,\phi), F_\theta(r,\theta,\phi), F_\phi(r,\theta,\phi)) \\ &= F_r(r,\theta,\phi)\mathbf{e}_r + F_\theta(r,\theta,\phi)\mathbf{e}_\theta + F_\phi(r,\theta,\phi)\mathbf{e}_\phi\end{aligned}$$

donde

$$\begin{aligned}F_r(r,\theta,\phi) &= \operatorname{sen}\phi\cos\theta\tilde{\mathbf{F}}_x(r,\theta,\phi) + \operatorname{sen}\theta\operatorname{sen}\phi\tilde{\mathbf{F}}_y(r,\theta,\phi) + \cos\phi\tilde{\mathbf{F}}_z(r,\theta,\phi) \\ F_\theta(r,\theta,\phi) &= -\operatorname{sen}\theta\tilde{\mathbf{F}}_x(r,\theta,\phi) + \cos\theta\tilde{\mathbf{F}}_y(r,\theta,\phi) \\ F_\phi(r,\theta,\phi) &= \cos\phi\cos\theta\tilde{\mathbf{F}}_x(r,\theta,\phi) + \operatorname{sen}\theta\cos\phi\tilde{\mathbf{F}}_y(r,\theta,\phi) - \operatorname{sen}\phi\tilde{\mathbf{F}}_z(r,\theta,\phi)\end{aligned}$$

en que a su vez

$$\tilde{\mathbf{F}}_{x(y,z)}(r,\theta,\phi) = F_{x(y,z)}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$$

Por otra parte, tenemos la expresión del campo rot $\mathbf{F}$ (en coordenadas cartesianas)

$$\operatorname{rot}\mathbf{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\mathbf{k}$$

y queremos ver entonces cómo se ve ésta en coordenadas esféricas cuando además en ella cambiamos el campo $\mathbf{F}$ por su versión en coordenadas esféricas $\tilde{\mathbf{F}}$. Sustituyendo en esta última expresión las variables x, y, z del sistema cartesiano por las variables r, θ, ϕ, según las ecuaciones de transformación $x = r\cos\theta\operatorname{sen}\phi$, $y = r\operatorname{sen}\theta\operatorname{sen}\phi$, $z = r\cos\phi$, nos quedaría la expresión de rot $\tilde{\mathbf{F}}(r,\theta,\phi)$ en la que, sustituyendo las expresiones de los vectores $\mathbf{i}, \mathbf{j}, \mathbf{k}$ del sistema ortonormal cartesiano en términos de los vectores $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi$ del sistema ortonormal esférico, es decir,

$$\begin{aligned}\mathbf{i} &= \operatorname{sen}\phi\cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta + \cos\phi\cos\theta\mathbf{e}_\phi \\ \mathbf{j} &= \operatorname{sen}\theta\operatorname{sen}\phi\mathbf{e}_r + \cos\theta\mathbf{e}_\theta + \operatorname{sen}\theta\cos\phi\mathbf{e}_\phi \\ \mathbf{k} &= \cos\phi\mathbf{e}_r - \operatorname{sen}\phi\mathbf{e}_\phi\end{aligned}$$

se obtendría finalmente al vector rot $\tilde{\mathbf{F}}(r,\theta,\phi)$ escrito como combinación lineal de los vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\phi$. De modo más explícito, se tendría una expresión del tipo

$$\begin{aligned}\operatorname{rot}\tilde{\mathbf{F}}(r,\theta,\phi) = &\left(\frac{\partial F_z}{\partial y}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right. \\ &\left.- \frac{\partial F_y}{\partial z}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right)(\operatorname{sen}\phi\cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta + \cos\phi\cos\theta\mathbf{e}_\phi) \\ &+ \left(\frac{\partial F_x}{\partial z}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right. \\ &\left.- \frac{\partial F_z}{\partial x}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right)(\operatorname{sen}\theta\operatorname{sen}\phi\mathbf{e}_r + \cos\theta\mathbf{e}_\theta + \operatorname{sen}\theta\cos\phi\mathbf{e}_\phi) \\ &+ \left(\frac{\partial F_y}{\partial x}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right. \\ &\left.- \frac{\partial F_x}{\partial y}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right)(\cos\phi\mathbf{e}_r - \operatorname{sen}\phi\mathbf{e}_\phi)\end{aligned}$$

En principio, esta es la expresión que buscamos. El problema es que las componentes del vector rot $\tilde{\mathbf{F}}(r, \theta, \phi)$ no están escritas en términos de las componentes del vector $\tilde{\mathbf{F}}(r, \theta, \phi)$. Y éste es precisamente el trabajo que comenzamos ahora. Se trata entonces de escribir a las derivadas parciales $\frac{\partial F_z}{\partial y}, \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z}, \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x}, \frac{\partial F_x}{\partial y}$ en términos de (las derivadas parciales de) las funciones componentes de $\tilde{\mathbf{F}}$, (es decir, en términos de F_r, F_θ y F_ϕ). Las expresiones

$$\begin{aligned} F_r(r, \theta, \phi) &= \operatorname{sen}\phi\cos\theta\tilde{F}_x(r, \theta, \phi) + \operatorname{sen}\theta\operatorname{sen}\phi\tilde{F}_y(r, \theta, \phi) + \cos\phi\tilde{F}_z(r, \theta, \phi) \\ F_\theta(r, \theta, \phi) &= -\operatorname{sen}\theta\tilde{F}_x(r, \theta, \phi) + \cos\theta\tilde{F}_y(r, \theta, \phi) \\ F_\phi(r, \theta, \phi) &= \cos\phi\cos\theta\tilde{F}_x(r, \theta, \phi) + \operatorname{sen}\theta\cos\phi\tilde{F}_y(r, \theta, \phi) - \operatorname{sen}\phi\tilde{F}_z(r, \theta, \phi) \end{aligned}$$

(donde $\tilde{F}_{x(y\,z)}(r, \theta, \phi) = F_{x(y\,z)}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$), se puede resolver para $\tilde{F}_x(r, \theta, \phi)$, $\tilde{F}_y(r, \theta, \phi)$ y $\tilde{F}_z(r, \theta, \phi)$, quedando

$$\begin{aligned} \tilde{F}_x(r, \theta, \phi) &= F_\phi(r, \theta, \phi)\cos\theta\cos\phi + F_r(r, \theta, \phi)\cos\theta\operatorname{sen}\phi - F_\theta(r, \theta, \phi)\operatorname{sen}\theta \\ \tilde{F}_y(r, \theta, \phi) &= F_\phi(r, \theta, \phi)\operatorname{sen}\theta\cos\phi + F_r(r, \theta, \phi)\operatorname{sen}\theta\operatorname{sen}\phi + F_\theta(r, \theta, \phi)\cos\theta \\ \tilde{F}_z(r, \theta, \phi) &= F_r(r, \theta, \phi)\cos\phi - F_\phi(r, \theta, \phi)\operatorname{sen}\phi \end{aligned}$$

Es decir, se tiene

$$\begin{aligned} F_x(x, y, z) &= F_\phi(r, \theta, \phi)\cos\theta\cos\phi + F_r(r, \theta, \phi)\cos\theta\operatorname{sen}\phi - F_\theta(r, \theta, \phi)\operatorname{sen}\theta \\ F_y(x, y, z) &= F_\phi(r, \theta, \phi)\operatorname{sen}\theta\cos\phi + F_r(r, \theta, \phi)\operatorname{sen}\theta\operatorname{sen}\phi + F_\theta(r, \theta, \phi)\cos\theta \\ F_z(x, y, z) &= F_r(r, \theta, \phi)\cos\phi - F_\phi(r, \theta, \phi)\operatorname{sen}\phi \end{aligned}$$

donde $x = r\cos\theta\operatorname{sen}\phi$, $y = r\operatorname{sen}\theta\operatorname{sen}\phi$, $z = r\cos\phi$. De estas expresiones podemos obtener las derivadas deseadas de las funciones F_x, F_y y F_z, tomando en cuenta que r, θ y ϕ dependen a su vez de x, y y z.

Antes de comenzar a derivar estas expresiones obsérvese que en este proceso aparecerán (según la regla de la cadena) las derivadas de r, θ y ϕ respecto de x, y y z. No interesan las expresiones que definen a r, θ y ϕ en términos de x, y y z, aunque si las tuviéramos podríamos obtener las derivadas mencionadas directamente de ellas. Las expresiones que tenemos son las que relacionan a x, y, z en términos de r, θ y ϕ (a saber, $x = r\cos\theta\operatorname{sen}\phi$, $y = r\operatorname{sen}\theta\operatorname{sen}\phi$, $z = r\cos\phi$). Así podemos usar el Teorema de la Función Inversa para obtener las derivadas de r, θ y ϕ respecto de x, y y z. En efecto, la función $\mathbf{H}: \mathbb{R}^+ \times [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$ dada por

$$\mathbf{H}(r, \theta, \phi) = (x, y, z) = (r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$$

tiene por jacobiano a

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = \det\mathbf{H}'(r, \theta, \phi) = -r^2\operatorname{sen}\phi$$

La función inversa $\mathbf{H}^{-1}(x, y, z) = (r, \theta, \phi)$ tiene por matriz jacobiana la inversa de la matriz jacobiana de $\mathbf{H}(r, \theta, \phi)$. Es decir,

$$
\begin{aligned}
(\mathbf{H}^{-1})'(x, y, z) &= (\mathbf{H}'(r, \theta, \phi))^{-1} \\
&= \begin{bmatrix} \cos\theta \operatorname{sen}\phi & -r \operatorname{sen}\theta \operatorname{sen}\phi & r\cos\theta\cos\phi \\ \operatorname{sen}\theta \operatorname{sen}\phi & r\cos\theta\cos\phi & r\operatorname{sen}\theta\cos\phi \\ \cos\phi & 0 & -r\operatorname{sen}\phi \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \cos\theta \operatorname{sen}\phi & \operatorname{sen}\theta \operatorname{sen}\phi & \cos\phi \\ -\dfrac{\operatorname{sen}\theta}{r\operatorname{sen}\phi} & \dfrac{\cos\theta}{r\operatorname{sen}\phi} & 0 \\ \dfrac{\cos\theta\cos\phi}{r} & \dfrac{\operatorname{sen}\theta\cos\phi}{r} & -\dfrac{\operatorname{sen}\phi}{r} \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{\partial r}{\partial x} & \dfrac{\partial r}{\partial y} & \dfrac{\partial r}{\partial z} \\ \dfrac{\partial \theta}{\partial x} & \dfrac{\partial \theta}{\partial y} & \dfrac{\partial \theta}{\partial z} \\ \dfrac{\partial \phi}{\partial x} & \dfrac{\partial \phi}{\partial y} & \dfrac{\partial \phi}{\partial z} \end{bmatrix}
\end{aligned}
$$

(las derivadas evaluadas en $(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$). Comencemos entonces a derivar la (primera) expresión

$$F_x(x, y, z) = F_\phi(r, \theta, \phi)\cos\theta\cos\phi + F_r(r, \theta, \phi)\cos\theta\operatorname{sen}\phi - F_\theta(r, \theta, \phi)\operatorname{sen}\theta$$

respecto de la variable y. Se tiene

$$
\begin{aligned}
\frac{\partial F_x}{\partial y} &= F_\phi \cos\theta(-\operatorname{sen}\phi)\frac{\partial\phi}{\partial y} + F_\phi(-\operatorname{sen}\theta)\frac{\partial\theta}{\partial y}\cos\phi \\
&\quad + \cos\theta\cos\phi\left(\frac{\partial F_\phi}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial F_\phi}{\partial \theta}\frac{\partial \theta}{\partial y} + \frac{\partial F_\phi}{\partial \phi}\frac{\partial \phi}{\partial y}\right) \\
&\quad + F_r\cos\theta\cos\phi\frac{\partial\phi}{\partial y} + F_r(-\operatorname{sen}\theta)\frac{\partial\theta}{\partial y}\operatorname{sen}\phi + \cos\theta\operatorname{sen}\phi\left(\frac{\partial F_r}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial F_r}{\partial \theta}\frac{\partial \theta}{\partial y} + \frac{\partial F_r}{\partial \phi}\frac{\partial \phi}{\partial y}\right) \\
&\quad - F_\theta\cos\theta\frac{\partial\theta}{\partial y} - \operatorname{sen}\theta\left(\frac{\partial F_\theta}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial F_\theta}{\partial \theta}\frac{\partial \theta}{\partial y} + \frac{\partial F_\theta}{\partial \phi}\frac{\partial \phi}{\partial y}\right)
\end{aligned}
$$

donde $\frac{\partial F_x}{\partial y}$ se evalúa en $(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$, F_r, F_θ, F_ϕ y sus derivadas están evaludas en (r, θ, ϕ) y las derivadas de r, θ y ϕ respecto de x, y, z están evaluadas en $(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)$. Sustituyendo y simplificando estas últimas derivadas se obtiene que

$$
\begin{aligned}
\frac{\partial F_x}{\partial y} &= -\frac{\operatorname{sen}\phi\cos\phi\operatorname{sen}\theta\cos\theta}{r}F_\phi - \frac{\operatorname{sen}\theta\cos\theta\cos\phi}{r\operatorname{sen}\phi}F_\phi + \operatorname{sen}\theta\cos\theta\operatorname{sen}\phi\cos\phi\frac{\partial F_\phi}{\partial r} \\
&\quad + \frac{\cos^2\theta\cos\phi}{r\operatorname{sen}\phi}\frac{\partial F_\phi}{\partial\theta} + \frac{\operatorname{sen}\theta\cos\theta\cos^2\phi}{r}\frac{\partial F_\phi}{\partial\phi} + \frac{\operatorname{sen}\theta\cos\theta\cos^2\phi}{r}F_r - \frac{\operatorname{sen}\theta\cos\theta}{r}F_r \\
&\quad + \operatorname{sen}\theta\cos\theta\operatorname{sen}^2\phi\frac{\partial F_r}{\partial r} + \frac{\cos^2\theta}{r}\frac{\partial F_r}{\partial\theta} + \frac{\operatorname{sen}\phi\cos\phi\operatorname{sen}\theta\cos\theta}{r}\frac{\partial F_r}{\partial\phi} - \frac{\cos^2\theta}{r\operatorname{sen}\phi}F_\theta \\
&\quad - \operatorname{sen}^2\theta\operatorname{sen}\phi\frac{\partial F_\theta}{\partial r} - \frac{\operatorname{sen}\theta\cos\theta}{r\operatorname{sen}\phi}\frac{\partial F_\theta}{\partial\theta} - \frac{\operatorname{sen}^2\theta\cos\phi}{r}\frac{\partial F_\theta}{\partial\phi}
\end{aligned}
$$

De esta misma manera, se obtienen las derivadas

$$\frac{\partial F_x}{\partial z} = \frac{\operatorname{sen}^2\phi\cos\theta}{r}F_\phi + \cos\theta\cos^2\phi\frac{\partial F_\phi}{\partial r} - \frac{\cos\theta\operatorname{sen}\phi\cos\phi}{r}\frac{\partial F_\phi}{\partial\phi} - \frac{\cos\theta\cos\phi\operatorname{sen}\phi}{r}F_r$$
$$+ \cos\theta\operatorname{sen}\phi\cos\phi\frac{\partial F_r}{\partial r} - \frac{\cos\theta\operatorname{sen}^2\phi}{r}\frac{\partial F_r}{\partial\phi} - \operatorname{sen}\theta\cos\phi\frac{\partial F_\theta}{\partial r} + \frac{\operatorname{sen}\theta\operatorname{sen}\phi}{r}\frac{\partial F_\theta}{\partial\phi}$$

$$\frac{\partial F_y}{\partial x} = -\frac{\operatorname{sen}\phi\cos\phi\operatorname{sen}\theta\cos\theta}{r}F_\phi - \frac{\operatorname{sen}\theta\cos\theta\cos\phi}{r\operatorname{sen}\phi}F_\phi + \operatorname{sen}\theta\cos\theta\operatorname{sen}\phi\cos\phi\frac{\partial F_\phi}{\partial r}$$
$$- \frac{\operatorname{sen}^2\theta\cos\phi}{r\operatorname{sen}\phi}\frac{\partial F_\phi}{\partial\theta} + \frac{\operatorname{sen}\theta\cos\theta\cos^2\phi}{r}\frac{\partial F_\phi}{\partial\phi} + \frac{\operatorname{sen}\theta\cos\theta\cos^2\phi}{r}F_r - \frac{\operatorname{sen}\theta\cos\theta}{r}F_r$$
$$+ \operatorname{sen}\theta\cos\theta\operatorname{sen}^2\phi\frac{\partial F_r}{\partial r} - \frac{\operatorname{sen}^2\theta}{r}\frac{\partial F_r}{\partial\theta} + \frac{\operatorname{sen}\phi\cos\phi\operatorname{sen}\theta\cos\theta}{r}\frac{\partial F_r}{\partial\phi} + \frac{\operatorname{sen}^2\theta}{r\operatorname{sen}\phi}F_\theta$$
$$+ \cos^2\theta\operatorname{sen}\phi\frac{\partial F_\theta}{\partial r} - \frac{\operatorname{sen}\theta\cos\theta}{r\operatorname{sen}\phi}\frac{\partial F_\theta}{\partial\theta} + \frac{\cos^2\theta\cos\phi}{r}\frac{\partial F_\theta}{\partial\phi}$$

$$\frac{\partial F_y}{\partial z} = \frac{\operatorname{sen}^2\phi\operatorname{sen}\theta}{r}F_\phi + \operatorname{sen}\theta\cos^2\phi\frac{\partial F_\phi}{\partial r} - \frac{\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi}{r}\frac{\partial F_\phi}{\partial\phi} - \frac{\operatorname{sen}\theta\cos\phi\operatorname{sen}\phi}{r}F_r$$
$$+ \operatorname{sen}\theta\operatorname{sen}\phi\cos\phi\frac{\partial F_r}{\partial r} - \frac{\operatorname{sen}\theta\operatorname{sen}^2\phi}{r}\frac{\partial F_r}{\partial\phi} + \cos\theta\cos\phi\frac{\partial F_\theta}{\partial r} - \frac{\cos\theta\operatorname{sen}\phi}{r}\frac{\partial F_\theta}{\partial\phi}$$

$$\frac{\partial F_z}{\partial x} = -\frac{\cos\theta\operatorname{sen}\phi\cos\phi}{r}F_r + \cos\theta\operatorname{sen}\phi\cos\phi\frac{\partial F_r}{\partial r} - \frac{\operatorname{sen}\theta\cos\phi}{r\operatorname{sen}\phi}\frac{\partial F_r}{\partial\theta} + \frac{\cos\theta\cos^2\phi}{r}\frac{\partial F_r}{\partial\phi}$$
$$- \frac{\cos\theta\cos^2\phi}{r}F_\phi - \cos\theta\operatorname{sen}^2\phi\frac{\partial F_\phi}{\partial r} + \frac{\operatorname{sen}\theta}{r}\frac{\partial F_\phi}{\partial\theta} - \frac{\cos\theta\operatorname{sen}\phi\cos\phi}{r}\frac{\partial F_\phi}{\partial\phi}$$

$$\frac{\partial F_z}{\partial y} = -\frac{\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi}{r}F_r + \operatorname{sen}\theta\operatorname{sen}\phi\cos\phi\frac{\partial F_r}{\partial r} + \frac{\cos\theta\cos\phi}{r\operatorname{sen}\phi}\frac{\partial F_r}{\partial\theta} + \frac{\operatorname{sen}\theta\cos^2\phi}{r}\frac{\partial F_r}{\partial\phi}$$
$$- \frac{\operatorname{sen}\theta\cos^2\phi}{r}F_\phi - \operatorname{sen}\theta\operatorname{sen}^2\phi\frac{\partial F_\phi}{\partial r} - \frac{\cos\theta}{r}\frac{\partial F_\phi}{\partial\theta} - \frac{\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi}{r}\frac{\partial F_\phi}{\partial\phi}$$

Sustituyendo entonces estas expresiones en

$$\operatorname{rot}\tilde{\mathbf{F}}(r,\theta,\phi) = \left(\frac{\partial F_z}{\partial y}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right.$$
$$\left. - \frac{\partial F_y}{\partial z}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right)(\operatorname{sen}\phi\cos\theta\mathbf{e}_r - \operatorname{sen}\theta\mathbf{e}_\theta + \cos\phi\cos\theta\mathbf{e}_\phi)$$
$$+ \left(\frac{\partial F_x}{\partial z}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right.$$
$$\left. - \frac{\partial F_z}{\partial x}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right)(\operatorname{sen}\theta\operatorname{sen}\phi\mathbf{e}_r + \cos\theta\mathbf{e}_\theta + \operatorname{sen}\theta\cos\phi\mathbf{e}_\phi)$$
$$+ \left(\frac{\partial F_y}{\partial x}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right.$$
$$\left. - \frac{\partial F_x}{\partial y}(r\cos\theta\operatorname{sen}\phi, r\operatorname{sen}\theta\operatorname{sen}\phi, r\cos\phi)\right)(\cos\phi\mathbf{e}_r - \operatorname{sen}\phi\mathbf{e}_\phi)$$

y simplificando, se obtiene finalmente la expresión que procurábamos del rotacional en coordenadas esféricas

$$\operatorname{rot} \tilde{\mathbf{F}}(r,\theta,\phi) = \left[\frac{1}{r \operatorname{sen}\phi}\frac{\partial}{\partial\phi}(\operatorname{sen}\phi F_\theta) - \frac{1}{r\operatorname{sen}\phi}\frac{\partial F_\phi}{\partial\theta}\right]\mathbf{e}_r + \left[\frac{1}{r\operatorname{sen}\phi}\frac{\partial F_r}{\partial\theta} - \frac{1}{r}\frac{\partial}{\partial r}(rF_\theta)\right]\mathbf{e}_\phi + \left[\frac{1}{r}\frac{\partial}{\partial r}(rF_\phi) - \frac{1}{r}\frac{\partial F_r}{\partial\phi}\right]\mathbf{e}_\theta$$

Ejercicios (Capítulo 9, Sección 4)

En los ejercicios 1—5, determine el rotacional del campo dado.

1. $\mathbf{F}(x, y, z) = (x, y, z)$.
2. $\mathbf{F}(x, y, z) = (yz, xz, xy)$.
3. $\mathbf{F}(x, y, z) = (2x + y, x - y + 3z, xyz)$.
4. $\mathbf{F}(x, y, z) = (x^2y, x + z^2, 3xz)$.
5. $\mathbf{F}(x, y, z) = (xe^xy, e^{-y}, z^3)$.

En los ejercicios 6—10, determine el rotacional del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ dado en coordenadas cilíndricas.

6. $\mathbf{F}(r, \theta, z) = \mathbf{e}_r + \mathbf{e}_\theta + \mathbf{e}_z$
7. $\mathbf{F}(r, \theta, z) = r\mathbf{e}_r + \theta\mathbf{e}_\theta + z\mathbf{e}_z$
8. $\mathbf{F}(r, \theta, z) = r\cos\theta\mathbf{e}_r + sen\theta\mathbf{e}_\theta + z^2\mathbf{e}_z$
9. $\mathbf{F}(r, \theta, z) = r^2\mathbf{e}_r + \cos\theta\mathbf{e}_\theta + \operatorname{sen}\theta\mathbf{e}_z$
10. $\mathbf{F}(r, \theta, z) = \cos^2\theta\mathbf{e}_r + r^3\mathbf{e}_\theta + \operatorname{sen}^2\theta\mathbf{e}_z$

En los ejercicios 11—15, determine el rotacional del campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ dado en coordenadas esféricas.

11. $\mathbf{F}(r, \theta, \phi) = \mathbf{e}_r + \mathbf{e}_\theta + \mathbf{e}_\phi$
12. $\mathbf{F}(r, \theta, \phi) = r\mathbf{e}_r + \phi\mathbf{e}_\theta + \theta\mathbf{e}_\phi$
13. $\mathbf{F}(r, \theta, \phi) = r\operatorname{sen}\theta\mathbf{e}_r + r\cos\phi\mathbf{e}_\theta$
14. $\mathbf{F}(r, \theta, \phi) = r\cos\theta\cos\phi\mathbf{e}_\theta + \operatorname{sen}\theta\cos\phi\mathbf{e}_\phi$
15. $\mathbf{F}(r, \theta, \phi) = (r^2 + \cos\theta\operatorname{sen}\phi)\mathbf{e}_r + 2r\operatorname{sen}\phi\mathbf{e}_\phi$

Compruebe que los campos dados en los ejercicios 16 y 17 son irrotacionales.

17. $\mathbf{F}(x, y, z) = (2xy^3z, 3x^2y^2z, x^2y^3)$. (Ver ejercicio 17 de la sección 4 del capítulo 7).

18. $\mathbf{F}(x, y, z) = (e^{x+2z}, xe^{x+2z}, 2xe^{x+2z})$. (Ver ejercicio 18 de la sección 4 del capítulo 7).

19. Demuestre que un campo constante $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (a, b, c)$, es irrotacional

20. Considere la función $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = 5x^2y + 2y^2 + 3xz^2$. Compruebe que $\operatorname{rot} \operatorname{grad} f = 0$.

21. Demuestre que el campo $\mathbf{F}: \mathbb{R}^3 - \{(0, 0)\} \to \mathbb{R}^2$ dado por $\mathbf{F}(x, y) = \frac{e^x}{x^2+y^2}(x \operatorname{sen} y - y \cos y, x \cos y + y \operatorname{sen} y, 0)$ es irrotacional, pero no conservativo. (Ver ejercicios 9 y 10 de la sección 10 del capítulo 7).

22. Considere el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$,

$$\mathbf{F}(x, y, z) = (3x^2 + 4y, 2x^2y + 5z, 3x + 4y + 5z^3)$$

a. Obtenga que $\operatorname{rot} \mathbf{F}(x, y, z) = (-1, -3, 4xy - 4)$.

b. Sea $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (r \cos t, r \operatorname{sen} t, 0)$. La imagen de $\boldsymbol{\lambda}$ es entonces un círculo en el plano xy, con centro en el origen, formando con el vector $N_p = (0, 0, 1)$ un "recorrido de mano derecha", lo cual significa que siguiendo el recorrido del círculo con el dedo índice de la mano derecha, y con el dedo pulgar perpendicular al índice, la dirección del pulgar será la del vector N_p. Calcule la integral de línea del campo $\mathbf{F}$ a lo largo de $\boldsymbol{\lambda}$.

c. Divida el resultado del inciso anterior entre el área del círculo, imagen de $\boldsymbol{\lambda}$. Compruebe que este resultado es el mismo al hacer el producto punto del vector $\operatorname{rot} \mathbf{F}(0, 0, 0)$ con el vector $N_p = (0, 0, 1)$.

d. Repita y compruebe los cálculos de los dos incisos anteriores considerando un círculo en el plano xz que forme un recorrido de mano derecha con el vector $N_p = (0, 1, 0)$.

e. Repita y compruebe los cálculos de los incisos b) y c) considerando un círculo en el plano yz que forme un recorrido de mano derecha con el vector $N_p = (1, 0, 0)$.

9.5 Teorema de Stokes

En esta sección estudiaremos otro de los resultados clásicos del cálculo en $\mathbb{R}^n$: el teorema de Stokes. Este teorema es una generalización a $\mathbb{R}^3$ del teorema de Green (sección 9, capítulo 7), el cual trabaja con campos en $\mathbb{R}^2$ y establece la relación entre la integral de línea del campo a lo largo de una curva cerrada simple, frontera de una región S en $\mathbb{R}^2$ (que llamamos "compacta") y una integral doble sobre la región S de "cierta" función (obviamente comprometida con el campo). El teorema de Stokes generalizará esta situación en el sentido siguiente: dejemos que la región S de $\mathbb{R}^2$ sea llevada (por una función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$) al espacio $\mathbb{R}^3$ y que su imagen sea una superficie simple, digamos K, cuya frontera es ∂K (imagen bajo $\mathbf{f}$ de la frontera de S). El teorema que estudiaremos establece la relación entre la integral de línea del campo $\mathbf{F}$ a lo largo de la frontera ∂K y la integral de superficie de "cierto" campo sobre la superficie K.

Sea K una superficie simple, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ (la cual, como sabemos, es inyectiva, de clase $\mathscr{C}^1$ y tal que $\mathbf{N_f} = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v}$ es nunca nulo en todo S). Consideramos

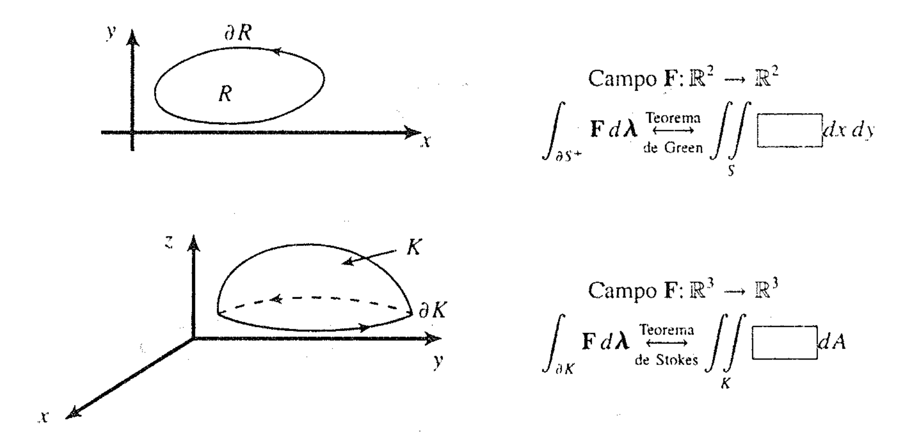


Figura 1. Relación entre el Teorema de Green y el de Stokes.

la superficie K junto con la orientación que esta parametrización le proporciona. La región S es una región del tipo I y II con frontera ∂S. Escribamos ∂S^+ para denotar la frontera ∂S junto con su orientación positiva. Supongamos que ∂S^+ es parametrizada por el camino $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$ seccionalmente $\mathscr{C}^1$. La imagen bajo $\mathbf{f}$ de ∂S^+ produce la frontera de K, la cual diremos que está "positivamente orientada", y la escribiremos como $\partial K^+ (= \mathbf{f}(\partial S^+))$. Esta es entonces una curva en $\mathbb{R}^3$. De hecho, la composición $\boldsymbol{\mu} = \mathbf{f} \circ \boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$ es un camino seccionalmente $\mathscr{C}^1$ que parametriza ∂K^+.

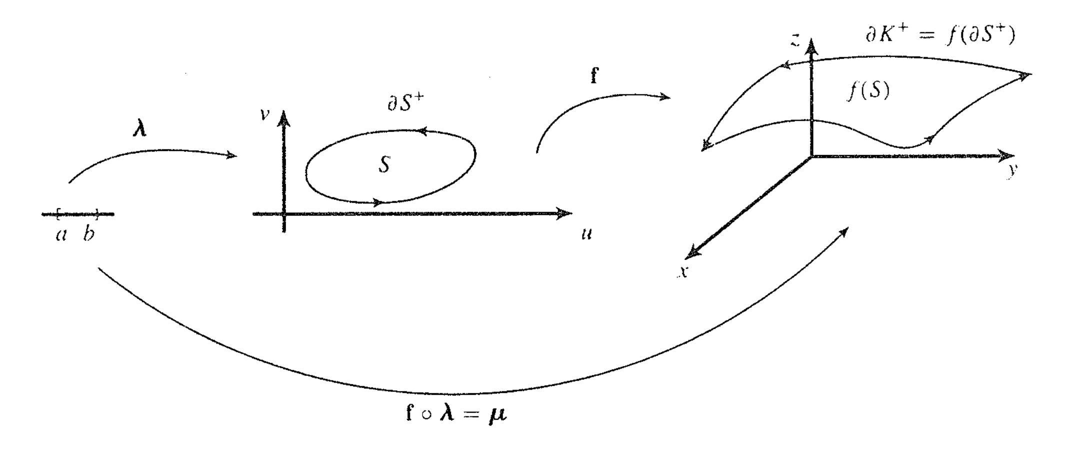


Figura 2. Las parametrizaciones de ∂S^+ y ∂K^+

Para la demostración del teorema de Stokes es necesario que la función $\mathbf{f}$ que parametriza a la superficie K sea una función de clase $\mathscr{C}^2$ (en tal caso se dice que $\mathbf{f}$ es una parametrización de clase $\mathscr{C}^2$ de K). Insistimos además en que la orientación dada de K es la que proporciona la parametrización $\mathbf{f}$ (es en este sentido en el que decimos que $\partial K^+ = \mathbf{f}(\partial S^+)$ es la frontera positivamente orientada de K).

Ahora estamos en posibilidades de enunciar y demostrar el teorema de Stokes.

Teorema 9.4.1 (Teorema de Stokes). Sea K una superficie simple orientable, parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ de clase $\mathscr{C}^2$, la cual proporciona la orientación de K, y sea $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ un campo vectorial de clase $\mathscr{C}^1$ definido en el conjunto abierto U de $\mathbb{R}^3$ que contiene a K. Entonces

$$\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} = \iint_{K_f} \operatorname{rot} \mathbf{F}\, dA$$

Demostración. Se trata solamente de hacer algunas operaciones (un poco largas, pero operaciones al fin, en las cuales se hará uso intenso de la regla de la cadena). Pongámonos de acuerdo con la notación: escribiremos

1. $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (x(u, v), y(u, v), z(u, v))$, donde entonces $x, y, z: S \subset \mathbb{R}^2 \to \mathbb{R}$ son funciones de clase $\mathscr{C}^2$.
2. $\boldsymbol{\lambda}: [a, b] \to \mathbb{R}^2$, $\boldsymbol{\lambda}(t) = (\lambda_1(t), \lambda_2(t))$ el camino seccionalmente $\mathscr{C}^1$ que parametriza ∂S^+.
3. $\boldsymbol{\mu} = \mathbf{f} \circ \boldsymbol{\lambda}: [a, b] \to \mathbb{R}^3$, $\boldsymbol{\mu}(t) = (\mu_1(t), \mu_2(t), \mu_3(t))$ el camino seccionalmente $\mathscr{C}^1$ que parametriza ∂K^+. Entonces $\mu_1(t) = x(\boldsymbol{\lambda}(t))$, $\mu_2(t) = y(\boldsymbol{\lambda}(t))$, $\mu_3(t) = z(\boldsymbol{\lambda}(t))$.
4. $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z))$, el campo vectorial de clase $\mathscr{C}^1$ considerado en el teorema.

Calculemos la integral de línea $\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu}$. Se tiene

$$\begin{aligned}
\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} &= \int_a^b \left[F_1(\mu(t))\mu'(t) + F_2(\mu(t))\mu'(t) + F_3(\mu(t))\mu'(t)\right] dt \\
&= \int_a^b \left[F_1(\mu(t))\frac{d}{dt}(x(\boldsymbol{\lambda}(t))) + F_2(\mu(t))\frac{d}{dt}(y(\boldsymbol{\lambda}(t))) + F_3(\mu(t))\frac{d}{dt}(z(\boldsymbol{\lambda}(t)))\right] dt \\
&= \int_a^b \left[F_1(\mu(t))\left(\frac{\partial x}{\partial u}(\boldsymbol{\lambda}(t))\boldsymbol{\lambda}_1'(t) + \frac{\partial x}{\partial v}(\boldsymbol{\lambda}(t))\boldsymbol{\lambda}_2'(t)\right)\right. \\
&\quad + F_2(\mu(t))\left(\frac{\partial y}{\partial u}(\boldsymbol{\lambda}(t))\boldsymbol{\lambda}_1'(t) + \frac{\partial y}{\partial v}(\boldsymbol{\lambda}(t))\boldsymbol{\lambda}_2'(t)\right) \\
&\quad \left. + F_3(\mu(t))\left(\frac{\partial z}{\partial u}(\boldsymbol{\lambda}(t))\boldsymbol{\lambda}_1'(t) + \frac{\partial z}{\partial v}(\boldsymbol{\lambda}(t))\boldsymbol{\lambda}_2'(t)\right)\right] dt \\
&= \int_a^b \left\{\left[F_1(f(\boldsymbol{\lambda}(t)))\frac{\partial x}{\partial u}(\boldsymbol{\lambda}(t)) + F_2(f(\boldsymbol{\lambda}(t)))\frac{\partial y}{\partial u}(\boldsymbol{\lambda}(t)) + F_3(f(\boldsymbol{\lambda}(t)))\frac{\partial z}{\partial u}(\boldsymbol{\lambda}(t))\right]\boldsymbol{\lambda}_1'(t)\right. \\
&\quad \left. + \left[F_1(f(\boldsymbol{\lambda}(t)))\frac{\partial x}{\partial v}(\boldsymbol{\lambda}(t)) + F_2(f(\boldsymbol{\lambda}(t)))\frac{\partial y}{\partial v}(\boldsymbol{\lambda}(t)) + F_3(f(\boldsymbol{\lambda}(t)))\frac{\partial z}{\partial v}(\boldsymbol{\lambda}(t))\right]\boldsymbol{\lambda}_2'(t)\right\} dt
\end{aligned}$$

Sea $\mathbf{G}: V \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{G} = (G_1, G_2)$, el campo vectorial definido en el conjunto abierto V de $\mathbb{R}^2$ que contiene a la región S, cuyas funciones coordenadas son

$$\begin{aligned}
G_1(u, v) &= F_1(\mathbf{f}(u, v))\frac{\partial x}{\partial u}(u, v) + F_2(\mathbf{f}(u, v))\frac{\partial y}{\partial u}(u, v) + F_3(\mathbf{f}(u, v))\frac{\partial z}{\partial u}(u, v) \\
G_2(u, v) &= F_1(\mathbf{f}(u, v))\frac{\partial x}{\partial v}(u, v) + F_2(\mathbf{f}(u, v))\frac{\partial y}{\partial v}(u, v) + F_3(\mathbf{f}(u, v))\frac{\partial z}{\partial v}(u, v)
\end{aligned}$$

Nótese que $\mathbf{G}$ es un campo de clase $\mathscr{C}^1$. La integral de línea de $\mathbf{G}$ a lo largo de ∂S^+ es

$$\begin{aligned}\int_{\partial S^+} \mathbf{G}\, d\boldsymbol{\lambda} &= \int_a^b \big(G_1(\boldsymbol{\lambda}(t))\lambda_1'(t) + G_2(\boldsymbol{\lambda}(t))\lambda_2'(t)\big)dt \\ &= \int_a^b \left\{ \left[F_1(\mathbf{f}(\boldsymbol{\lambda}(t)))\frac{\partial x}{\partial u}(\boldsymbol{\lambda}(t)) + F_2(\mathbf{f}(\boldsymbol{\lambda}(t)))\frac{\partial y}{\partial u}(\boldsymbol{\lambda}(t)) + F_3(\mathbf{f}(\boldsymbol{\lambda}(t)))\frac{\partial z}{\partial u}(\boldsymbol{\lambda}(t))\right]\lambda_1'(t)\right. \\ &\quad \left. + \left[F_1(\mathbf{f}(\boldsymbol{\lambda}(t)))\frac{\partial x}{\partial v}(\boldsymbol{\lambda}(t)) + F_2(\mathbf{f}(\boldsymbol{\lambda}(t)))\frac{\partial y}{\partial v}(\boldsymbol{\lambda}(t)) + F_3(\mathbf{f}(\boldsymbol{\lambda}(t)))\frac{\partial z}{\partial v}(\boldsymbol{\lambda}(t))\right]\lambda_2'(t)\right\}dt\end{aligned}$$

la cual es igual a la última expresión que se había establecido para la integral del línea $\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu}$. Tenemos entonces que

$$\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} = \int_{\partial S^+} \mathbf{G} \cdot d\boldsymbol{\lambda} = \xrightarrow{\text{Teorema de Green}} = \iint_S \left(\frac{\partial G_2}{\partial u} - \frac{\partial G_1}{\partial v}\right) du\, dv$$

Calculemos las derivadas parciales de las funciones coordenadas de $\mathbf{G}$ indicadas en la última expresión.

$$\begin{aligned}\frac{\partial G_2}{\partial u} &= \frac{\partial}{\partial u}\left[F_1(\mathbf{f}(u,v))\frac{\partial x}{\partial v}(u,v) + F_2(\mathbf{f}(u,v))\frac{\partial y}{\partial v}(u,v) + F_3(\mathbf{f}(u,v))\frac{\partial z}{\partial v}(u,v)\right] \\ &= F_1(\mathbf{f}(u,v))\frac{\partial^2 x}{\partial u \partial v}(u,v) \\ &\quad + \frac{\partial x}{\partial v}(u,v)\left[\frac{\partial F_1}{\partial x}(\mathbf{f}(u,v))\frac{\partial x}{\partial u}(u,v) + \frac{\partial F_1}{\partial y}(\mathbf{f}(u,v))\frac{\partial y}{\partial u}(u,v) + \frac{\partial F_1}{\partial z}(\mathbf{f}(u,v))\frac{\partial z}{\partial u}(u,v)\right] \\ &\quad + F_2(\mathbf{f}(u,v))\frac{\partial^2 y}{\partial u \partial v}(u,v) \\ &\quad + \frac{\partial y}{\partial v}(u,v)\left[\frac{\partial F_2}{\partial x}(\mathbf{f}(u,v))\frac{\partial x}{\partial u}(u,v) + \frac{\partial F_2}{\partial y}(\mathbf{f}(u,v))\frac{\partial y}{\partial u}(u,v) + \frac{\partial F_2}{\partial z}(\mathbf{f}(u,v))\frac{\partial z}{\partial u}(u,v)\right] \\ &\quad + F_3(\mathbf{f}(u,v))\frac{\partial^2 z}{\partial u \partial v}(u,v) \\ &\quad + \frac{\partial z}{\partial v}(u,v)\left[\frac{\partial F_3}{\partial x}(\mathbf{f}(u,v))\frac{\partial x}{\partial u}(u,v) + \frac{\partial F_3}{\partial y}(\mathbf{f}(u,v))\frac{\partial y}{\partial u}(u,v) + \frac{\partial F_3}{\partial z}(\mathbf{f}(u,v))\frac{\partial z}{\partial u}(u,v)\right]\end{aligned}$$

y

$$\begin{aligned}\frac{\partial G_1}{\partial v} &= \frac{\partial}{\partial v}\left[F_1(\mathbf{f}(u,v))\frac{\partial x}{\partial u}(u,v) + F_2(\mathbf{f}(u,v))\frac{\partial y}{\partial u}(u,v) + F_3(\mathbf{f}(u,v))\frac{\partial z}{\partial u}(u,v)\right] \\ &= F_1(\mathbf{f}(u,v))\frac{\partial^2 x}{\partial v \partial u}(u,v) \\ &\quad + \frac{\partial x}{\partial u}(u,v)\left[\frac{\partial F_1}{\partial x}(\mathbf{f}(u,v))\frac{\partial x}{\partial v}(u,v) + \frac{\partial F_1}{\partial y}(\mathbf{f}(u,v))\frac{\partial y}{\partial v}(u,v) + \frac{\partial F_1}{\partial z}(\mathbf{f}(u,v))\frac{\partial z}{\partial v}(u,v)\right] \\ &\quad + F_2(\mathbf{f}(u,v))\frac{\partial^2 y}{\partial v \partial u}(u,v)\end{aligned}$$

$$+ \frac{\partial y}{\partial u}(u, v)\left[\frac{\partial F_2}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial v}(u, v) + \frac{\partial F_2}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial v}(u, v) + \frac{\partial F_2}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial v}(u, v)\right]$$
$$+ F_3(\mathbf{f}(u, v))\frac{\partial^2 z}{\partial v \partial u}(u, v)$$
$$+ \frac{\partial z}{\partial u}(u, v)\left[\frac{\partial F_3}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial v}(u, v) + \frac{\partial F_3}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial v}(u, v) + \frac{\partial F_3}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial v}(u, v)\right]$$

de modo que

$$\begin{aligned}
\frac{\partial G_2}{\partial u} - \frac{\partial G_1}{\partial v} &= F_1(\mathbf{f}(u, v))\left[\frac{\partial^2 x}{\partial u \partial v}(u, v) - \frac{\partial^2 x}{\partial v \partial u}(u, v)\right] \\
&+ F_2(\mathbf{f}(u, v))\left[\frac{\partial^2 y}{\partial u \partial v}(u, v) - \frac{\partial^2 y}{\partial v \partial u}(u, v)\right] + F_3(\mathbf{f}(u, v))\left[\frac{\partial^2 z}{\partial u \partial v}(u, v) - \frac{\partial^2 z}{\partial v \partial u}(u, v)\right] \\
&+ \frac{\partial x}{\partial v}(u, v)\left[\frac{\partial F_1}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial u}(u, v) + \frac{\partial F_1}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial u}(u, v) + \frac{\partial F_1}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial u}(u, v)\right] \\
&+ \frac{\partial y}{\partial v}(u, v)\left[\frac{\partial F_2}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial u}(u, v) + \frac{\partial F_2}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial u}(u, v) + \frac{\partial F_2}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial u}(u, v)\right] \\
&+ \frac{\partial z}{\partial v}(u, v)\left[\frac{\partial F_3}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial u}(u, v) + \frac{\partial F_3}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial u}(u, v) + \frac{\partial F_3}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial u}(u, v)\right] \\
&- \frac{\partial x}{\partial u}(u, v)\left[\frac{\partial F_1}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial v}(u, v) + \frac{\partial F_1}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial v}(u, v) + \frac{\partial F_1}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial v}(u, v)\right] \\
&- \frac{\partial y}{\partial u}(u, v)\left[\frac{\partial F_2}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial v}(u, v) + \frac{\partial F_2}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial v}(u, v) + \frac{\partial F_2}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial v}(u, v)\right] \\
&- \frac{\partial z}{\partial u}(u, v)\left[\frac{\partial F_3}{\partial x}(\mathbf{f}(u, v))\frac{\partial x}{\partial v}(u, v) + \frac{\partial F_3}{\partial y}(\mathbf{f}(u, v))\frac{\partial y}{\partial v}(u, v) + \frac{\partial F_3}{\partial z}(\mathbf{f}(u, v))\frac{\partial z}{\partial v}(u, v)\right]
\end{aligned}$$

Partiendo de que las derivadas parciales de segundo orden cruzadas de las funciones $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$ son iguales (pues se está asumiendo clase $\mathscr{C}^2$ para estas funciones), la expresión anterior queda como

$$\begin{aligned}
&\frac{\partial F_1}{\partial y}(\mathbf{f}(u, v))\left[\frac{\partial x}{\partial v}(u, v)\frac{\partial y}{\partial u}(u, v) - \frac{\partial x}{\partial u}(u, v)\frac{\partial y}{\partial v}(u, v)\right] \\
&+ \frac{\partial F_1}{\partial z}(\mathbf{f}(u, v))\left[\frac{\partial x}{\partial v}(u, v)\frac{\partial z}{\partial u}(u, v) - \frac{\partial x}{\partial u}(u, v)\frac{\partial z}{\partial v}(u, v)\right] \\
&+ \frac{\partial F_2}{\partial x}(\mathbf{f}(u, v))\left[\frac{\partial x}{\partial u}(u, v)\frac{\partial y}{\partial v}(u, v) - \frac{\partial x}{\partial v}(u, v)\frac{\partial y}{\partial u}(u, v)\right] \\
&+ \frac{\partial F_2}{\partial z}(\mathbf{f}(u, v))\left[\frac{\partial y}{\partial v}(u, v)\frac{\partial z}{\partial u}(u, v) - \frac{\partial y}{\partial u}(u, v)\frac{\partial z}{\partial v}(u, v)\right] \\
&+ \frac{\partial F_3}{\partial x}(\mathbf{f}(u, v))\left[\frac{\partial x}{\partial u}(u, v)\frac{\partial z}{\partial v}(u, v) - \frac{\partial x}{\partial v}(u, v)\frac{\partial z}{\partial u}(u, v)\right] \\
&+ \frac{\partial F_3}{\partial y}(\mathbf{f}(u, v))\left[\frac{\partial y}{\partial u}(u, v)\frac{\partial z}{\partial v}(u, v) - \frac{\partial y}{\partial v}(u, v)\frac{\partial z}{\partial u}(u, v)\right]
\end{aligned}$$

Nótese que el contenido de los corchetes en la expresión anterior son, a menos de signos, los jacobianos de las funciones x y y, x y z, o y y z, respecto de u y v. Podemos entonces reescribir esta última expresión en términos de estos jacobianos como

$$\left(\frac{\partial F_3}{\partial y}(\mathbf{f}(u, v)) - \frac{\partial F_2}{\partial z}(\mathbf{f}(u, v))\right)\frac{\partial(y, z)}{\partial(u, v)}$$
$$+ \left(\frac{\partial F_1}{\partial z}(\mathbf{f}(u, v)) - \frac{\partial F_3}{\partial x}(\mathbf{f}(u, v))\right)\frac{\partial(z, x)}{\partial(u, v)}$$
$$+ \left(\frac{\partial F_2}{\partial x}(\mathbf{f}(u, v)) - \frac{\partial F_1}{\partial y}(\mathbf{f}(u, v))\right)\frac{\partial(x, y)}{\partial(u, v)}$$

expresión que no es más que el producto punto del vector

$$\operatorname{rot} \mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

evaluado en $\mathbf{f}(u, v)$, con el vector

$$\mathbf{N_f} = \left(\frac{\partial(y, z)}{\partial(u, v)}, \frac{\partial(z, x)}{\partial(u, v)}, \frac{\partial(x, y)}{\partial(u, v)}\right)$$

evaluado en (u, v). Juntando todas las piezas, tenemos entonces que la integral de línea del campo $\mathbf{F}$ a lo largo de la frontera de K positivamente orientada es

$$\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} = \iint_S \left(\frac{\partial G_2}{\partial u} - \frac{\partial G_1}{\partial v}\right) du\, dv = \iint_S \operatorname{rot} \mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N_f}(u, v) du\, dv = \iint_{K_f} \operatorname{rot} \mathbf{F}\, dA$$

como se quería demostrar. Q.E.D.

Como habíamos dicho, el teorema de Stokes generaliza al teorema de Green, quedando este último como un caso particular del primero. En efecto, consideremos el campo $\tilde{\mathbf{F}}: \{(x, y, z) \mid (x, y) \in U, z = 0\} \to \mathbb{R}^3$, donde U es un conjunto abierto de $\mathbb{R}^2$, dado por

$$\tilde{\mathbf{F}}(x, y, 0) = (\tilde{F}_1(x, y, 0), \tilde{F}_2(x, y, 0), 0)$$

Es claro que podemos identificar este campo con el campo $\mathbf{F}: U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{F} = (F_1, F_2)$, $\mathbf{F}(x, y) = \tilde{\mathbf{F}}(x, y, 0)$ (y entonces $F_1(x, y) = \tilde{F}_1(x, y, 0)$, $F_2(x, y) = \tilde{F}_2(x, y, 0)$). El rotacional del campo $\tilde{\mathbf{F}}$ es

$$\operatorname{rot} \mathbf{F} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \tilde{F}_1 & \tilde{F}_2 & 0 \end{bmatrix} = \left(\frac{\partial \tilde{F}_2}{\partial z}, \frac{\partial \tilde{F}_1}{\partial z}, \frac{\partial \tilde{F}_2}{\partial x} - \frac{\partial \tilde{F}_1}{\partial y}\right) = \left(0, 0 \frac{\partial \tilde{F}_2}{\partial x} - \frac{\partial \tilde{F}_1}{\partial y}\right)$$

Consideremos la superficie simple K parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, definida en la región S de $\mathbb{R}^2$ la cual está contenida en U, $\mathbf{f}(u, v) = (u, v, 0)$. Obsérvese que K no es más que una copia en el plano xy de $\mathbb{R}^3$ de la región S de $\mathbb{R}^2$ (conservando esta copia la orientación positiva de

S). Como $\mathbf{N_f}(u, v) = \frac{\partial \mathbf{f}}{\partial u} \times \frac{\partial \mathbf{f}}{\partial v} = (1, 0, 0) \times (0, 1, 0) = (0, 0, 1)$, según el teorema de Stokes tenemos que

$$\begin{aligned}\int_{\partial K^+} \tilde{\mathbf{F}} \cdot d\mu &= \iint_{K_f} \operatorname{rot} \tilde{\mathbf{F}}\, dA = \iint_S \operatorname{rot} \mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N_f}(u, v) du\, dv \\ &= \iint_S \left(0, 0, \frac{\partial \tilde{F}_2}{\partial x}(u, v, 0) - \frac{\partial \tilde{F}_1}{\partial y}(u, v, 0)\right) \cdot (0, 0, 1) du\, dv \\ &= \iint_S \left(\frac{\partial \tilde{F}_2}{\partial x}(u, v, 0) - \frac{\partial \tilde{F}_1}{\partial y}(u, v, 0)\right) du\, dv\end{aligned}$$

Haciendo las identificaciones correspondientes del campo $\tilde{\mathbf{F}}$ con el campo $\mathbf{F}$ y de la región S en el plano uv, con la superficie K en el plano xy, (junto con la del camino $\boldsymbol{\mu}$ que parametriza ∂K^+ y el camino $\boldsymbol{\lambda}$ que parametriza a ∂S^+), tenemos que

$$\int_{\partial S^+} \mathbf{F} \cdot d\boldsymbol{\lambda} = \iint_S \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dx\, dy$$

que es la fórmula del teorema de Green.

Veamos algunos ejemplos.

Ejemplo 1. Sea $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ el campo $\mathbf{F}(x, y, z) = (xy, 2xz, 3yz)$ y sea K la superficie correspondiente a la porción del plano $x + y + z = 1$ que se encuentra en el primer octante, con sus normales apuntando hacia "afuera" (hacia el primer octante). Verifiquemos el teorema de Stokes en este caso. Una parametrización de K está dada por $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (u, v, 1 - u - v)$, donde $S = \{(u, v)|0 \leq u \leq 1, 1 \leq v \leq 1 - u\}$. Obsérvese que $\mathbf{N_f} = (1, 1, 1)$, de modo que esta parametrización produce la orientación de K. Un camino $\boldsymbol{\lambda}$ que parametriza ∂S^+ es $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3$, en el que

$$\begin{aligned}&\boldsymbol{\lambda}_1: [0, 1] \to \mathbb{R}^2, \boldsymbol{\lambda}_1(t) = (t, 0) \\ &\boldsymbol{\lambda}_2: [0, 1] \to \mathbb{R}^2, \boldsymbol{\lambda}_2(t) = (1 - t, t) \\ &\boldsymbol{\lambda}_3: [0, 1] \to \mathbb{R}^2, \boldsymbol{\lambda}_3(t) = (0, 1 - t)\end{aligned}$$

de modo que un camino $\boldsymbol{\mu}$ que parametriza a ∂K^+ es $\boldsymbol{\mu} = \mathbf{f} \circ \boldsymbol{\lambda}$, $\boldsymbol{\mu} = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2 + \boldsymbol{\mu}_3$. Es decir

$$\begin{aligned}&\boldsymbol{\mu}_1: [0, 1] \to \mathbb{R}^3, \boldsymbol{\mu}_1(t) = f(\boldsymbol{\lambda}_1(t)) = f(t, 0) = (t, 0, 1 - t) \\ &\boldsymbol{\mu}_2: [0, 1] \to \mathbb{R}^3, \boldsymbol{\mu}_2(t) = f(\boldsymbol{\lambda}_2(t)) = f(1 - t, t) = (1 - t, t, 0) \\ &\boldsymbol{\mu}_3: [0, 1] \to \mathbb{R}^3, \boldsymbol{\mu}_3(t) = f(\boldsymbol{\lambda}_3(t)) = f(0, 1 - t) = (0, 1 - t, t)\end{aligned}$$

Calculemos la integral de línea del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\mu}$. Se tiene

$$\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} = \int_{\boldsymbol{\mu}_1} \mathbf{F} \cdot d\boldsymbol{\mu}_1 + \int_{\boldsymbol{\mu}_2} \mathbf{F} \cdot d\boldsymbol{\mu}_2 + \int_{\boldsymbol{\mu}_3} \mathbf{F} \cdot d\boldsymbol{\mu}_3$$

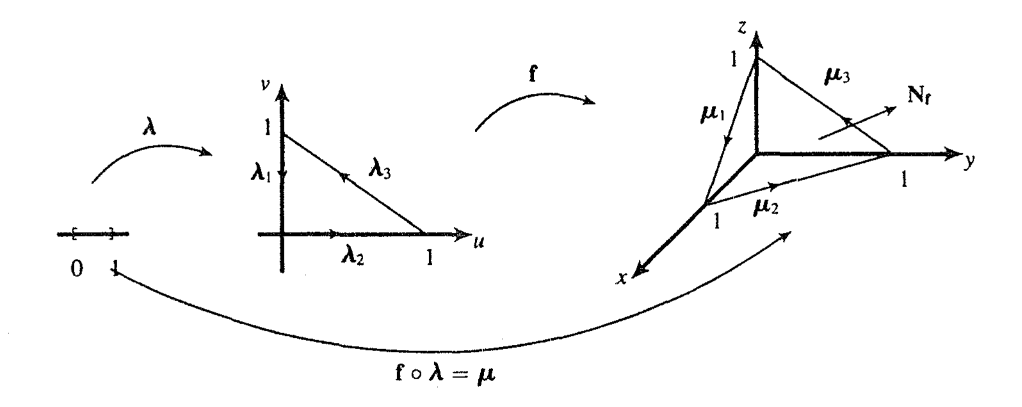


Figura 3. Las parametrizaciones de los caminos del ejemplo 1

donde

$$\int_{\mu_1} \mathbf{F}\cdot d\boldsymbol{\mu}_1 = \int_0^1 \mathbf{F}(\boldsymbol{\mu}_1(t))\cdot \boldsymbol{\mu}_1'(t)dt = \int_0^1 (0, 2t(1-t), 0)\cdot(1, 0, -1)dt = 0$$

$$\int_{\mu_2} \mathbf{F}\cdot d\boldsymbol{\mu}_2 = \int_0^1 \mathbf{F}(\boldsymbol{\mu}_2(t))\cdot \boldsymbol{\mu}_2'(t)dt = \int_0^1 ((1-t)t, 0, 0)\cdot(-1, 1, 0)dt = \int_0^1 (t^2-t)dt = -1/6$$

$$\int_{\mu_3} \mathbf{F}\cdot d\boldsymbol{\mu}_3 = \int_0^1 \mathbf{F}(\boldsymbol{\mu}_3(t))\cdot \boldsymbol{\mu}_3'(t)dt = \int_0^1 (0, 0, 3(1-t)t)\cdot(0, -1, 1)dt = \int_0^1 3(1-t)tdt = 1/2$$

de modo que

$$\int_{\partial K^+} \mathbf{F}\cdot d\boldsymbol{\mu} = -\frac{1}{6}+\frac{1}{2} = \frac{1}{3}$$

Calculemos ahora la integral de superficie del campo rot $\mathbf{F}$ sobre K. Se tiene

$$\operatorname{rot}\mathbf{F} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ xy & 2xz & 3yz \end{bmatrix} = (3z-2x, 0, 2z-x)$$

Entonces

$$\iint_{K_f} \operatorname{rot}\mathbf{F}\, dA = \iint_S \operatorname{rot}\mathbf{F}(\mathbf{f}(u,v))\cdot \mathbf{N_f}(u,v)du\, dv$$

$$= \iint_S \big(3(1-u-v)-2u, 0, 2(1-u-v)-u\big)\cdot(1,1,1)du\, dv$$

$$= \int_0^1\int_0^{1-u} (5-8u-5v)du\, dv = 1/3$$

Es decir, se tiene

$$\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} = \frac{1}{3} = \iint_{K_f} \operatorname{rot} \mathbf{F}\, dA \qquad \blacksquare$$

El teorema de Stokes puede aplicarse a superficies seccionalmente simples, es decir, superficies parametrizadas por funciones $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$ (de clase $\mathscr{C}^2$), definidas en regiones S que se pueden descomponer en subregiones del tipo (I) y (II).

Ejemplo 2. Consideremos el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$ dado por $\mathbf{F}(x, y, z) = (x^2y, 3x^3z, yz^2)$ y sea K el cilindro $\{(x, y, z)|x^2 + y^2 = 1, -1 \le z \le 1\}$, con sus normales apuntando hacia dentro de él. Esta es una superficie seccionalmente simple parametrizada por la función $\mathbf{f}: S \subset \mathbb{R}^2 \to \mathbb{R}^3$,

$$\mathbf{f}(u, v) = \left(\frac{u}{\sqrt{u^2 + v^2}}, \frac{v}{\sqrt{u^2 + v^2}}, \sqrt{u^2 + v^2} - 2 \right)$$

(la cual es de clase $\mathscr{C}^2$ y produce la orientación establecida para K), definida en el anillo circular

$$S = \{(u, v)|1 \le u^2 + v^2 \le 9\}$$

(ver ejemplo 6 de la sección 1, capítulo 8). Verifiquemos el teorema de Stokes con este campo y esta superficie. La frontera de S consta de los círculos $C_1: x^2 + y^2 = 9$, y $C_2: x^2 + y^2 = 1$. La orientación positiva de S se logra recorriendo C_1 en dirección antihoraria y C_2 en dirección horaria. Entonces un camino $\boldsymbol{\lambda}$ que parametriza a ∂S^+ es $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2$, en el que

$$\boldsymbol{\lambda}_1: [0, 2\pi] \to \mathbb{R}^2, \boldsymbol{\lambda}_1(t) = (3\cos t, 3\operatorname{sen} t)$$
$$\boldsymbol{\lambda}_2: [0, 2\pi] \to \mathbb{R}^2, \boldsymbol{\lambda}_2(t) = (\cos t, -\operatorname{sen} t)$$

Entonces la orientación positiva de la frontera de K queda parametrizada por el camino $\boldsymbol{\mu} = \mathbf{f} \circ \boldsymbol{\lambda} = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2$, en el que

$$\boldsymbol{\mu}_1: [0, 2\pi] \to \mathbb{R}^3, \boldsymbol{\mu}_1(t) = \mathbf{f}(\boldsymbol{\lambda}_1(t)) = \mathbf{f}(3\cos t, 3\operatorname{sen} t) = (\cos t, \operatorname{sen} t, 1)$$
$$\boldsymbol{\mu}_2: [0, 2\pi] \to \mathbb{R}^3, \boldsymbol{\mu}_2(t) = \mathbf{f}(\boldsymbol{\lambda}_2(t)) = \mathbf{f}(\cos t, -\operatorname{sen} t) = (\cos t, -\operatorname{sen} t, -1)$$

Calculemos la integral de línea del campo $\mathbf{F}$ a lo largo de ∂K^+. Se tiene

$$\int_{\partial K^+} \mathbf{F} \cdot d\boldsymbol{\mu} = \int_{\boldsymbol{\mu}_1} \mathbf{F} \cdot d\boldsymbol{\mu}_1 + \int_{\boldsymbol{\mu}_2} \mathbf{F} \cdot d\boldsymbol{\mu}_2$$

donde

$$\begin{aligned} \int_{\boldsymbol{\mu}_1} \mathbf{F} \cdot d\boldsymbol{\mu}_1 &= \int_0^{2\pi} \mathbf{F}(\boldsymbol{\mu}_1(t)) \cdot \boldsymbol{\mu}_1'(t)dt \\ &= \int_0^{2\pi} (\cos^2 t \operatorname{sen} t, 3\cos^3 t, \operatorname{sen} t) \cdot (-\operatorname{sen} t, \cos t, 0)dt \\ &= \int_0^{2\pi} (-\cos^2 t \operatorname{sen}^2 t + 3\cos^4 t)dt = 2\pi \end{aligned}$$

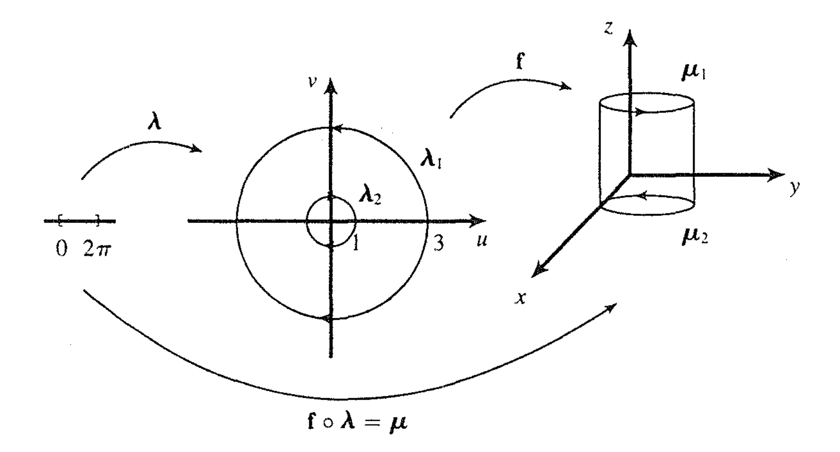


Figura 4. Las parametrizaciones de los caminos del ejemplo 2.

y

$$\int_{\mu_2} \mathbf{F} \cdot d\mu_1 = \int_0^{2\pi} \mathbf{F}(\mu_2(t)) \cdot \mu_2'(t)dt$$

$$= \int_0^{2\pi} (-\cos^2 t \operatorname{sen} t, -3\cos^3 t, -\operatorname{sen} t) \cdot (-\operatorname{sen} t, -\cos t, 0)\, dt$$

$$= \int_0^{2\pi} (\cos^2 t \operatorname{sen}^2 t + 3\cos^4 t)dt = 5\pi/2$$

Entonces

$$\int_{\partial K^+} \mathbf{F} \cdot d\mu = 2\pi + 5\pi/2 = 9\pi/2$$

Calculemos ahora la integral de superficie del campo rot $\mathbf{F}$ sobre K. Se tiene

$$\operatorname{rot} \mathbf{F} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ x^2 y & 3x^3 z & yz^2 \end{bmatrix} = \left(z^2 - 3x^3, 0, x^2(9z - 1)\right)$$

Puesto que

$$\mathbf{N_f}(u, v) = \left(-\frac{u}{u^2 + v^2}, -\frac{v}{u^2 + v^2}, 0\right)$$

(ver ejemplo 6 de la sección 6, capítulo 8), se tiene

$$\begin{aligned}
\iint\limits_{K_f} \operatorname{rot} \mathbf{F}\, dA &= \iint\limits_{S} \operatorname{rot} \mathbf{F}(\mathbf{f}(u, v)) \cdot \mathbf{N}_{\mathbf{f}}(u, v)\, du\, dv \\
&= \iint\limits_{S} \left[\left(\sqrt{u^2+v^2}\right)^2 - 3\left(\frac{u}{\sqrt{u^2+v^2}}\right)^3, 0, \left(\frac{u}{\sqrt{u^2+v^2}}\right)^3\left(9\sqrt{u^2+v^2} - 19\right)\right] \cdot \\
&\qquad \left[-\frac{u}{u^2+v^2}, -\frac{v}{u^2+v^2}, 0\right] du\, dv \\
&= \iint\limits_{S} \frac{u}{u^2+v^2}\left[-\left(\sqrt{u^2+v^2} - 2\right)^2 + 3\left(\frac{u}{\sqrt{u^2+v^2}}\right)^3\right] du\, dv \\
&\overset{\text{polares}}{\longrightarrow} \int_0^{2\pi} \int_1^3 \frac{\cos\theta}{r}\left(3\cos^3\theta - (r-2)^2\right) r\, dr\, d\theta = 9\pi/2
\end{aligned}$$

resultado que verifica el teorema de Stokes en este ejemplo. ■

Con la ayuda del teorema de Stokes podemos ver que el rotacional de un campo en $\mathbb{R}^3$ está relacionado con la "circulación del campo por unidad de área", al igual que ocurre con la rotación de un campo en $\mathbb{R}^2$ (sección 10 del capítulo 7). En efecto, sea $\mathbf{F}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ un campo de clase $\mathscr{C}^1$ definido en el conjunto abierto U de $\mathbb{R}^3$. Sea $\mathbf{p} \in U$ y considere un disco S_ϵ con centro en $\mathbf{p}$ y radio $\epsilon > 0$. Sea $\mathbf{N}$ un vector ortogonal al (plano en que se encuentra el) disco en el punto $\mathbf{p}$ (figura 5).

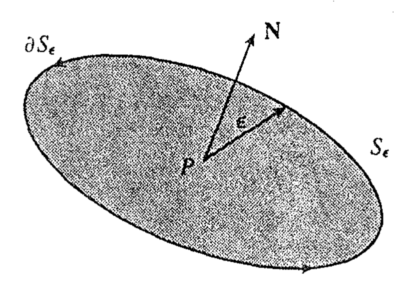


Figura 5. El disco alrededor del punto **p**.

La circulación del campo $\mathbf{F}$ alrededor de la frontera ∂S_ϵ (positivamente orientada) es, como se definió en la sección 10 del capítulo 7, la integral de línea de $\mathbf{F}$ a lo largo de ∂S_ϵ. Es decir

$$\text{Circulación de } \mathbf{F} \text{ alrededor de } \partial S_\epsilon = \int_{\partial S_\epsilon} \mathbf{F} \cdot d\boldsymbol{\mu}$$

(donde $\boldsymbol{\mu}$ es un camino que parametriza la frontera positivamente orientada de S_ϵ). El teorema de Stokes establece que

$$\int_{\partial S_\epsilon} \mathbf{F} \cdot d\boldsymbol{\mu} = \iint\limits_{S_\epsilon} \operatorname{rot} \mathbf{F} \cdot \mathbf{N}\, dA$$

Aplicando el teorema del valor medio a la integral de superficie de esta expresión (ver sección 1.1), existe un punto $\tilde{\mathbf{p}} \in S_\epsilon$ que

$$\iint_{S_\epsilon} \operatorname{rot} \mathbf{F} \cdot \mathbf{N}\, dA = \operatorname{rot} \mathbf{F}(\tilde{\mathbf{p}}) \cdot \mathbf{N}(\mathbf{p}) \iint_{S_\epsilon} dA = \operatorname{rot} \mathbf{F}(\tilde{\mathbf{p}}) \cdot \mathbf{N}(\mathbf{p}) \text{ Área de } S_\epsilon$$

Entonces

$$\begin{aligned} \operatorname{rot} \mathbf{F}(\tilde{\mathbf{p}}) \cdot \mathbf{N}(\mathbf{p}) &= \frac{1}{\text{Área de } S_\epsilon} \iint_{S_\epsilon} \operatorname{rot} \mathbf{F} \cdot \mathbf{N}\, dA = \frac{1}{\text{Área de } S_\epsilon} \int_{\partial S_\epsilon} \mathbf{F} \cdot d\boldsymbol{\mu} \\ &= \frac{\text{Circulación de } F \text{ alrededor de } S_\epsilon}{\text{Área de } S_\epsilon} \end{aligned}$$

Al hacer tender ϵ a cero, se obtiene, por la continuidad de las derivadas parciales de las funciones coordenadas de $\mathbf{F}$, que ("el punto $\tilde{\mathbf{p}}$ tiende al punto $\mathbf{p}$)

$$\begin{aligned} \operatorname{rot} \mathbf{F}(\mathbf{p}) \cdot \mathbf{N}(\mathbf{p}) &= \lim_{\epsilon \to 0} \operatorname{rot} \mathbf{F}(\tilde{\mathbf{p}}) \cdot \mathbf{N}(\mathbf{p}) \\ &= \lim_{\epsilon \to 0} \frac{\text{Circulación de } \mathbf{F} \text{ alrededor de } S_\epsilon}{\text{Área de } S_\epsilon} \end{aligned}$$

y así, la componente del vector rot $\mathbf{F}(\mathbf{p})$ en la dirección de la normal $\mathbf{N}$ es el límite de la circulación del campo $\mathbf{F}$ "alredor del punto $\mathbf{p}$", o bien, "alrededor de un recorrido que circunda a $\mathbf{p}$", por unidad de área, cuando este recorrido "tiende hacia el punto $\mathbf{p}$". Obsérvese que el hecho de que S_ϵ sea un *disco* es irrelevante. Esta misma idea vale para cualquier superficie (de las consideradas en el teorema de Stokes) que contenga a $\mathbf{p}$, y que se contraiga hacia este punto.

Ejercicios (Capítulo 9, Sección 5)

1. Verifique el teorema de Stokes con el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (2x, 3xy, z)$ y la superficie $K = \{(x, y, z) | x^2 + y^2 + z^2 = 1, z \geq 0\}$.

2. Verifique el teorema de Stokes con el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x+2y, 3x+z, x+y+z)$ y el cilindro $x^2 + y^2 = 4$, $-1 \leq z \leq 1$.

3. Verifique el teorema de Stokes con el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (z^2 + 1, 2z, 2xz + 2y)$ y la porción de esfera $x^2 + y^2 + z^2 = 1$ en el primer octante.

4. Considere el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x^2, xy, z^2)$. Sea C la intersección del cilindro $x^2 + y^2 = 1$ con el plano $x + y + z = 1$.

a. Demuestre que una parametrización de C está dada por $\boldsymbol{\lambda}: [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\lambda}(t) = (\cos t, \operatorname{sen} t, 1 - \cos t - \operatorname{sen} t)$.

b. Calcule la integral de línea del campo $\mathbf{F}$ a lo largo del camino $\boldsymbol{\lambda}$.

c. Considere la superficie K que consiste en la porción del plano $x + y + z = 1$ dentro de C. Calcule la integral de superficie del campo rot $\mathbf{F}$ sobre la superficie K. Constate que se satisfaga el teorema de Stokes.

9.6 Grad., Div., Rot. Las fórmulas clásicas del análisis vectorial

En el capítulo 2 se introdujo el concepto de gradiente de una función real diferenciable de n variables. En el capítulo 7 apareció el concepto de divergencia de un campo vectorial y , en este capítulo, el de rotacional de un campo en $\mathbb{R}^3$. Lo que haremos en esta sección será unir estas ideas y ver algunas propiedades que resultan al tratar de combinarlas. Las fórmulas resultantes que establecen tales propiedades son fórmulas muy importantes cuando se hace uso del llamado "Análisis Vectorial" (el cálculo que hemos hecho, pero con campos y/o funciones reales en $\mathbb{R}^3$) en el estudio de otras ramas de la matemática o de la física.

En toda esta sección se trabajará solamente en el espacio $\mathbb{R}^3$. Denotaremos con U a un conjunto abierto de $\mathbb{R}^3$. Comencemos por recordar las definiciones correspondientes. Si $f\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ es una función real diferenciable, se define el *gradiente de* f en el punto $\mathbf{p} \in U$, denotado por grad $f(\mathbf{p})$, como

$$\operatorname{grad} f(\mathbf{p}) = \left(\frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p}), \frac{\partial f}{\partial z}(\mathbf{p})\right)$$

Para el campo diferenciable $\mathbf{F}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$, se define su *divergencia* en el punto $\mathbf{p} \in U$, denotada por div $\mathbf{F}(\mathbf{p})$, como

$$\operatorname{div} \mathbf{F}(\mathbf{p}) = \frac{\partial F_1}{\partial x}(\mathbf{p}) + \frac{\partial F_2}{\partial y}(\mathbf{p}) + \frac{\partial F_3}{\partial z}(\mathbf{p})$$

y su *rotacional* en el punto $\mathbf{p} \in U$, denotado por rot $\mathbf{F}(\mathbf{p})$, como

$$\operatorname{rot} \mathbf{F}(\mathbf{p}) = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

(las derivadas evaluadas en $\mathbf{p}$).

Definimos el operador ∇ (nabla), como

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

Este es un operador que nos permite un manejo formal del gradiente, la divergencia y el rotacional. De hecho, nótese que si definimos la acción de ∇ sobre la función $f\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}$ como

$$\nabla f = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$$

tenemos que

$$\boxed{\nabla f = \operatorname{grad} f}$$

Viendo el operador ∇ como "un vector en $\mathbb{R}^3$", podemos tomar su producto punto con el campo $\mathbf{F}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$, quedando

$$\nabla \cdot \mathbf{F} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot (F_1, F_2, F_3) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

Es decir, se tiene

$$\boxed{\nabla \cdot \mathbf{F} = \operatorname{div} \mathbf{F}}$$

Haciendo ahora el producto cruz de ∇ con $\mathbf{F}$ se obtiene

$$\nabla \times \mathbf{F} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times (F_1, F_2, F_3) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{bmatrix}$$

y así:

$$\boxed{\nabla \times \mathbf{F} = \operatorname{rot} \mathbf{F}}$$

Viendo así al gradiente, la divergencia y el rotacional, es posible, por medio de un manejo puramente formal, establecer algunas propiedades de estas operaciones, como por ejemplo, su linealidad. En efecto, (dejamos como ejercicio la verificación de la linealidad para el gradiente) si $\mathbf{F}, \mathbf{G}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ son dos campos diferenciables y $\lambda \in \mathbb{R}$, se tiene

$$\operatorname{div}(\mathbf{F} + \lambda \mathbf{G}) = \nabla \cdot (\mathbf{F} + \lambda \mathbf{G}) = \nabla \cdot \mathbf{F} + \lambda \nabla \cdot \mathbf{G} = \operatorname{div} \mathbf{F} + \lambda \operatorname{div} \mathbf{G}$$
$$\operatorname{rot}(\mathbf{F} + \lambda \mathbf{G}) = \nabla \times (\mathbf{F} + \lambda \mathbf{G}) = \nabla \times \mathbf{F} + \lambda \nabla \times \mathbf{G} = \operatorname{rot} \mathbf{F} + \lambda \operatorname{rot} \mathbf{G}$$

donde se hizo uso del carácter lineal del producto punto y el producto cruz.

A continuación se establecen algunas de las fórmulas en las que se muestra el comportamiento de las operaciones grad, div y rot.

Si $f, g: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ son dos funciones diferenciables, $\mathbf{F}, \mathbf{G}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ dos campos diferenciables y $\lambda \in \mathbb{R}$, se tiene

1.	$\operatorname{grad}(f + \lambda g) = \operatorname{grad} f + \lambda \operatorname{grad} g$		1. $\nabla(f + \lambda g) = \nabla f + \lambda \nabla g$
2.	$\operatorname{div}(\mathbf{F} + \lambda \mathbf{G}) = \operatorname{div} \mathbf{F} + \lambda \operatorname{div} \mathbf{G}$		2. $\nabla \cdot (\mathbf{F} + \lambda \mathbf{G}) = \nabla \cdot \mathbf{F} + \lambda \nabla \cdot \mathbf{G}$
3.	$\operatorname{rot}(\mathbf{F} + \lambda \mathbf{G}) = \operatorname{rot} \mathbf{F} + \lambda \operatorname{rot} \mathbf{G}$		3. $\nabla \times (\mathbf{F} + \lambda \mathbf{G}) = \nabla \times \mathbf{F} + \lambda \nabla \times \mathbf{G}$
4.	$\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f$	o bien	4. $\nabla(fg) = f \nabla g + g \nabla f$
5.	$\operatorname{div}(\mathbf{F} \times \mathbf{G}) = (\operatorname{rot} \mathbf{F}) \cdot \mathbf{G} - (\operatorname{rot} \mathbf{G}) \cdot \mathbf{F}$		5. $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = (\nabla \times \mathbf{F}) \cdot \mathbf{G} - (\nabla \times \mathbf{G}) \cdot \mathbf{F}$
6.	$\operatorname{div}(f\mathbf{F}) = f \operatorname{div} \mathbf{F} + \mathbf{F} \cdot \operatorname{grad} f$		6. $\nabla \cdot (f\mathbf{F}) = f \nabla \cdot \mathbf{F} + \mathbf{F} \cdot \nabla f$
7.	$\operatorname{rot}(f\mathbf{F}) = f \operatorname{rot} \mathbf{F} + (\operatorname{grad} f) \times \mathbf{F}$		7. $\nabla \times (f\mathbf{F}) = f \nabla \times \mathbf{F} + (\nabla f) \times \mathbf{F}$

La verificación de fórmulas de este tipo consiste en operaciones de rutina que dejamos al lector. Presentamos solamente algunos pasos que muestran la validez de la fórmula 6. Se tiene

$$\begin{aligned} \operatorname{div}(f\mathbf{F}) &= \frac{\partial}{\partial x}(fF_1) + \frac{\partial}{\partial y}(fF_2) + \frac{\partial}{\partial z}(fF_3) \\ &= f\frac{\partial F_1}{\partial x} + F_1\frac{\partial f}{\partial x} + f\frac{\partial F_2}{\partial y} + F_2\frac{\partial f}{\partial y} + f\frac{\partial F_3}{\partial z} + F_3\frac{\partial f}{\partial z} \\ &= f\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot F + F \cdot \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) f = \\ &= f \operatorname{div} F + F \cdot \operatorname{grad} f \end{aligned}$$

Si asumimos clase $\mathscr{C}^2$ para las funciones f y g y el campo $\mathbf{F}$, se tienen además las fórmulas

8. rot grad $f = \mathbf{0}$ o bien 8. $\nabla \times (\nabla f) = \mathbf{0}$
9. div rot $\mathbf{F} = 0$ 9. $\nabla \cdot (\nabla \times F) = \mathbf{0}$
10. div(grad $f \times$ grad $g) = \mathbf{0}$ 10. $\nabla \cdot (\nabla f \times \nabla g) = \mathbf{0}$

La fórmula 8 ya quedó establecida en la sección 4. Veamos la validez de la fórmula 9. Se tiene

$$\begin{aligned}
\operatorname{div}(\operatorname{rot}\mathbf{F}) &= \operatorname{div}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) \\
&= \left(\frac{\partial^2 F_3}{\partial x \partial y} - \frac{\partial^2 F_3}{\partial y \partial x}\right) + \left(\frac{\partial^2 F_1}{\partial y \partial z} - \frac{\partial^2 F_1}{\partial z \partial y}\right) + \left(\frac{\partial^2 F_2}{\partial z \partial x} - \frac{\partial^2 F_2}{\partial x \partial z}\right) = 0
\end{aligned}$$

(pues las funciones F_1, F_2, F_3 son de clase $\mathscr{C}^2$ y entonces sus derivadas parciales de segundo orden mixtas son iguales).

Para ver la validez de la fórmula 10 podemos utilizar las fórmulas 5 y 8 de la manera siguiente

$$\nabla \cdot (\nabla f \times \nabla g) = (\nabla \times \nabla f) \cdot \nabla g - (\nabla \times \nabla g) \cdot \nabla f = 0 \cdot \nabla g - 0 \cdot \nabla f = 0$$

Sea $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ una función de clase $\mathscr{C}^2$. Consideremos la divergencia de su gradiente. Este es

$$\operatorname{div}(\operatorname{grad} f) = \operatorname{div}\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

o bien, en términos del operador ∇

$$\begin{aligned}
\nabla \cdot (\nabla f) &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \\
&= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}
\end{aligned}$$

Se define el *operador de Laplace*, o *laplaciano*, denotado por ∇^2, como

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

La acción del laplaciano sobre la función f (de clase $\mathscr{C}^2$) es

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

NOTA: tanto el operador ∇ como el operador ∇^2 se definen por lo general como:

$$\begin{aligned}
\nabla &= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right) \in \mathbb{R}^n \\
\nabla^2 &= \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2}
\end{aligned}$$

El operador ∇ actúa sobre la función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ como

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) = \text{grad } f$$

y el operador ∇^2 actúa sobre f (asumiendo clase $\mathscr{C}^2$ para esta función) como

$$\nabla^2 f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}$$ ■

Para terminar esta sección haremos algunos comentarios sobre un tipo de funciones llamadas armónicas y probaremos un resultado interesante acerca de ellas.

Se dice que la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, de clase $\mathscr{C}^2$, definida en el conjunto abierto U de $\mathbb{R}^n$, es *armónica* si satisface la ecuación de Laplace

$$\nabla^2 f = 0$$

Es decir, si

$$\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} = 0$$

En el ejemplo 4 de la sección 12 del capítulo 2 y en los ejercicios 55, 56, 57 de la sección 2 del capítulo 3, ya habían aparecido este tipo de funciones.

Las funciones armónicas $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ tienen una propiedad muy interesante: sea $\Omega \subset U$ una región en $\mathbb{R}^3$ (de las consideradas en el teorema de la divergencia) y $K = \partial\Omega$ la frontera de Ω (K es entonces una superficie cerrada en $\mathbb{R}^3$); supongamos que f se anula en K (es decir, $f(x, y, z) = 0$ $\forall(x, y, z) \in K$). Se puede demostrar (lo haremos) que de hecho f vale cero en *todo* Ω. En otras palabras, si una función armónica se anula en la frontera de la región $\Omega \subset \mathbb{R}^3$, es porque de hecho la función es la función cero en Ω.

La validez de esta afirmación será una consecuencia inmediata del inciso a) del siguiente teorema, en el que se muestran dos fórmulas clásicas en el análisis vectorial conocidas como "fórmulas de Green".

Teorema 9.6.1 Sean $f, g: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ dos funciones de clase $\mathscr{C}^2$ definidas en el conjunto abierto U de $\mathbb{R}^3$ y sea $\Omega \subset U$ una región en $\mathbb{R}^3$ (de las consideradas en el teorema de la divergencia). Supongamos que la superficie $K = \partial\Omega$ está orientada con sus normales apuntando hacia su exterior (digamos que ϕ es una parametrización de K). Entonces

a. $$\iint_{K_\phi} f\nabla g \, dA = \iiint_{\Omega} (f\nabla^2 g + \nabla f \cdot \nabla g)\, dx\, dy\, dz$$

b. $$\iint_{K_\phi} (f\nabla g - g\nabla f)\, dA = \iiint_{\Omega} (f\nabla^2 g - g\nabla^2 f)\, dx\, dy\, dz$$

Demostración. Se aplica el teorema de la divergencia

$$\iint_{K_\phi} \mathbf{F}\, dA = \iiint_{\Omega} \text{div}\, \mathbf{F}\, dx\, dy\, dz$$

a. Con el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = f\nabla g$, el teorema de la divergencia

$$\iint_{K_\phi} f\nabla g\, dA = \iiint_{\Omega} \nabla \cdot (f\nabla g)\, dx\, dy\, dz$$

pero (ver fórmula 6)

$$\nabla \cdot (f\nabla g) = f\nabla^2 g + \nabla f \cdot \nabla g$$

lo que prueba la validez de **a**

b. Con el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = f\nabla g - g\nabla f$, se tiene

$$\begin{aligned}\iint_{K_\phi} (f\nabla g - g\nabla f)\, dA &= \iiint_{\Omega} \nabla \cdot (f\nabla g - g\nabla f)\, dx\, dy\, dz \\ &= \iiint_{\Omega} (\nabla \cdot (f\nabla g) - \nabla \cdot (g\nabla f))\, dx\, dy\, dz \\ &= \iiint_{\Omega} (f\nabla^2 g + \nabla f \cdot \nabla g - g\nabla^2 f - \nabla g \cdot \nabla f)\, dx\, dy\, dz \\ &= \iiint_{\Omega} (f\nabla^2 g - g\nabla^2 f)\, dx\, dy\, dz \qquad \text{Q.E.D.}\end{aligned}$$

Supongamos entonces que $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ es una función armónica y que $f(x, y, z) = 0$ $\forall (x, y, z) \in K = \partial\Omega$.

Si en la fórmula del inciso *a.* del teorema anterior ponemos $g = f$, nos queda

$$\iint_{K_\phi} f\nabla f\, dA = \iiint_{\Omega} (f\nabla^2 f + \nabla f \cdot \nabla f)\, dx\, dy\, dz$$

Siendo f armónica ($\nabla^2 f = 0$), la expresión anterior se ve como

$$\iint_{K_\phi} f\nabla g\, dA = \iiint_{\Omega} \|\nabla f\|^2\, dx\, dy\, dz$$

Ya que la función f se anula en K, el integrando de la integral de superficie de la expresión anterior es cero. Esta integral vale entonces cero. Nos queda así que

$$\iiint_{\Omega} \|\nabla f\|^2\, dx\, dy\, dz = 0$$

La función $\tilde{f} = \|\nabla f\|^2$ es una función de clase $\mathscr{C}^1$ *no negativa en* Ω (i.e. $\tilde{f} \geq 0$ en Ω). Puesto que su integral en Ω vale 0, concluimos que

$$\|\nabla f\|^2 = 0 \qquad (\text{en } \Omega)$$

Es decir, $\left(\frac{\partial f}{\partial x}^2\right) + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 = 0$, de donde a su vez se concluye que

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0$$

Así pues, f es una función constante en Ω. Como f es continua y vale cero en $\partial\Omega$, concluimos finalmente que **f** debe también valer cero en toda la región Ω, como se quería.

Ejercicios (Capítulo 9, Sección 6)

1. Sean $f, g: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ dos funciones diferenciables, y λ un número real. Demuestre que:

a. $\nabla(f + \lambda g) = \nabla f + \lambda \nabla g$

b. $\nabla(fg) = f\nabla g + g\nabla f$

c. $\nabla \cdot (\nabla f \times \nabla g) = 0$ (suponga clase $\mathscr{C}^2$ para f y g).

2. Sean $\mathbf{F}, \mathbf{G}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ dos campos vectoriales diferenciables y $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable. Demuestre que:

a. $\nabla \cdot (\mathbf{F} + \lambda\mathbf{G}) = \nabla \cdot \mathbf{F} + \lambda\nabla \cdot \mathbf{G}$, $\lambda \in \mathbb{R}$.

b. $\nabla \times (\mathbf{F} + \lambda\mathbf{G}) = \nabla \times \mathbf{F} + \lambda\nabla \times \mathbf{G}$, $\lambda \in \mathbb{R}$.

c. $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = (\nabla \times \mathbf{F}) \cdot \mathbf{G} - (\nabla \times \mathbf{G}) \cdot \mathbf{F}$

d. $\nabla \times (f\mathbf{F}) = f\nabla \times \mathbf{F} + (\nabla f) \times \mathbf{F}$

3. Verifique el teorema 9.6.1 con las funciones $f, g: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^2 + y^2 + z^2$, $g(x, y, z) = xyz$, y Ω la región contenida en la esfera unitaria $x^2 + y^2 + z^2 = 1$.

Capítulo 10

Formas diferenciales

En la historia que se ha venido desarrollando en todo el libro, hasta el capítulo anterior, se vislumbra una de las características propias de la matemática: se trata de generar ideas y resultados cada vez más generales con los cuales se pueda lograr un mejor entendimiento de la teoría que se va dejando atrás.

Así por ejemplo, en la sección 3 del capítulo 3 se estudió la versión general de la regla de la cadena que muestra que la manera como se derivan funciones compuestas es, "en esencia", el mismo resultado que conocíamos ya del curso de cálculo de funciones de una sola variable, y que engloba también los casos particulares estudiados en la sección 2 de ese capítulo. Esta misma situación se presenta con el estudio de las funciones implícitas e inversas de las secciones 4, 5 y 6 del mismo capítulo.

Por otra parte, las integrales n-múltiples estudiadas en el capítulo 6 son una generalización natural, a dimensiones mayores, de las integrales "simples" del primer curso de cálculo. A este respecto, recordamos lo que nos dice uno de los matemáticos más importantes del siglo XX, David Hilbert: "... En respuesta, permítaseme señalar cómo internamente, por la misma naturaleza de las ciencias matemáticas, cualquier progreso real lleva dentro de sí el descubrimiento de herramientas más incisivas y métodos más simples, los cuales facilitan al mismo tiempo la comprensión de las teorías previas y eliminan los desarrollos más escabrosos...". Las generalizaciones que hemos comentado anteriormente son buenos ejemplos de las "herramientas más incisivas" señaladas por Hilbert. Así pues, en los nueve capítulos anteriores se han explorado las ideas del cálculo en dimensiones superiores, dejando atrás las ideas y resultados del cálculo de funciones de una sola variable. Como fruto de este estudio hemos obtenido resultados que son muy importantes por las conexiones que presentan entre las generalizaciones mencionadas. Los ejemplos más importantes de esto son:

1. El teorema de Green (capítulo 7, sección 9), cuya "fórmula" es

$$\int_{\partial S^+} \mathbf{F} \cdot d\lambda = \iint_S \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) dx\, dy$$

el cual establece la relación entre la integral de línea del campo $\mathbf{F} = (M, N)$ sobre la frontera de una cierta región S en $\mathbb{R}^2$, y la integral doble de "cierta función" sobre la región S misma.

2. El teorema de la divergencia (capítulo 9, sección 3), cuya "fórmula" es

$$\iint_{K_f} \mathbf{F}\, dA = \iiint_\Omega \operatorname{div} F\, dx\, dy\, dz$$

el cual establece la relación entre la integral de superficie de un campo F en $\mathbb{R}^3$, sobre una superficie cerrada K_f, y la integral triple de la divergencia de F sobre la región Ω que encierra K_f.

3. El teorema de Stokes (capítulo 9, sección 5), cuya "fórmula" es

$$\int_{\partial K^+} F \cdot d\mu = \iint_{K_f} \operatorname{rot} F \, dA$$

la cual establece la relación entre la integral de superficie del rotacional de un campo F en $\mathbb{R}^3$, sobre una superficie K_f con frontera positivamente orientada ∂K_f^+, y la integral de línea del campo F sobre ∂K_f^+.

Lo que queremos hacer en este capítulo es (además de terminar el libro) explorar la teoría que nos permita ver que estos tres teoremas son en realidad *un solo resultado*. Este será el ejemplo más vivo, en este libro, de lo que comentábamos en el primer párrafo: el resultado que presentamos en la sección 6 de este capítulo cubre, en cierto sentido, *TODO* el material de los nueve capítulos anteriores. Por supuesto que para lograr esto, necesitaremos explorar nuevas ideas y conceptos que nos permitan llegar al nivel necesario y tener una visión adecuada con la cual podremos resumir, prácticamente en una sola fórmula, el contenido de todo el libro.

Este capítulo está destinado para aquellos lectores que gusten de este tipo de aventuras. Es, definitivamente, el capítulo "más matemático" de todo el libro. En las próximas secciones se sentirá el peculiar sabor de las mieles de la abstracción matemática, con las cuales nos gustaría dejar al lector en calidad de "pasajero en tránsito", para que se pueda introducir posteriormente en el estudio de otras ramas de la matemática cuyos prerrequisitos estén relacionados con el cálculo en $\mathbb{R}^n$.

10.1 Definiciones preliminares. Suma y producto de formas

En esta sección presentamos los objetos matemáticos con los que se trabajará en el presente capítulo. La metodología a seguir, como lo sugiere una presentación introductoria y elemental del tema, será definiendo estos objetos como *expresiones formales* que cumplen ciertas reglas al operar con ellos. Así pues, el lector debe estar dispuesto a aceptar definiciones que contrastan mucho con las que han aparecido en los capítulos anteriores. Aunque esta manera de presentar a las formas diferenciales puede llegar a parecer mágica y artificial, la ganancia que se tiene a posteriori es inmensa en comparación con el posible mal sabor de boca que puede causar (a los lectores de mente matemática muy exigente) el manejar expresiones puramente formales.

A continuación nos referiremos a la expresión "forma diferencial", simplemente como "forma". Usaremos letras griegas como ω, η, τ para denotar las formas.

Sea $U \subseteq \mathbb{R}^n$ un conjunto abierto de $\mathbb{R}^n$. Una *0-forma* (cero forma) definida en U es una función diferenciable $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. Si la función es de clase $\mathscr{C}^k$, decimos que la 0-forma es de clase $\mathscr{C}^k$. Por ejemplo, la función $f: \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = xy^2 + z^3$ es una 0-forma de clase $\mathscr{C}^\infty$ (es decir, de clase $\mathscr{C}^k$ $\forall k \in \mathbb{N}$), definida en todo el espacio $\mathbb{R}^3$ (escribimos tal forma $\omega = xy^2 + z^3$). La función $g: U \subseteq \mathbb{R}^4 \to \mathbb{R}$, $g(x_1, x_2, x_3, x_4) = x_1 x_2 \ln(x_3) \ln(x_4)$ es una 0-forma de clase $\mathscr{C}^\infty$ definida en el conjunto abierto $U = \{(x_1, x_2, x_3, x_4) | x_3 > 0, x_4 > 0\} \subset \mathbb{R}^4$.

Una *1-forma* (uno forma) definida en $U \subseteq \mathbb{R}^n$ es una expresión del tipo

$$\omega = f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n$$

donde $f_1, f_2, \ldots, f_n: U \subseteq \mathbb{R}^n \to \mathbb{R}$ son funciones diferenciables. Si éstas son de clase $\mathscr{C}^k$ decimos que la 1-forma ω es de clase $\mathscr{C}^k$. En particular cada una de las expresiones dx_i, $i = 1, 2, \ldots, n$, es una 1-forma de clase $\mathscr{C}^\infty$ definida en $\mathbb{R}^n$ (para la que $f_i = 1$, $f_j = 0$ con $j \neq i$). Obsérvese que cada sumando $f_i dx_i$ en la expresión de ω se puede considerar como un "producto" de la 0-forma f_i por la 1-forma dx_i. Posteriormente definiremos con precisión el producto de formas. Por ejemplo, la expresión $\omega = (x + y)\,dx + xy\,dy$ es una 1-forma de clase $\mathscr{C}^\infty$ en $\mathbb{R}^2$.

Una *2-forma* (dos forma) definida en $U \subseteq \mathbb{R}^n$ es una expresión del tipo

$$\omega = \sum_{i,j}^{n} f_{ij}\, dx_i dx_j$$

donde $f_{ij}: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $1 \leq i, j \leq n$, son funciones diferenciables. Como en los casos anteriores, si estas funciones son de clase $\mathscr{C}^k$ decimos que ω es una 2-forma de clase $\mathscr{C}^k$.

Advertimos que las expresiones dx_i que han aparecido en las 1-formas y en las 2-formas son ... ¡simplemente expresiones! (así fueron definidas). Sin embargo, si al ver escrito dx, al lector le salta la idea de "la diferencial de x" (como la que aparecía en las integrales del primer curso de cálculo, o en el capítulo 6 con las integrales múltiples), no debe inquietarse. Al contrario, ésta es una manera conveniente de pensar estos objetos, pues ayuda a "entender" algunas de las operaciones entre formas que se estudiarán posteriormente. Las expresiones dx_i tienen un caracter anticonmutativo cuando se multiplican entre sí. De hecho, se define (la primera regla del juego)

$$dx_i dx_j = -dx_j dx_i$$

para todo i, j. Una consecuencia inmediata de este hecho es que

$$dx_i dx_i = 0$$

Es decir, el producto de una expresión dx_i por ella misma es igual a cero, pues, por la anticonmutatividad de su producto se tiene $dx_i dx_i = -dx_i dx_i$, de donde $dx_i dx_i = 0$.

Aplicando esta propiedad podemos escribir la definición de una 2-forma en una expresión "más simple" (con menos sumandos). En efecto, considere, por ejemplo, el caso $n = 3$. En principio, una 2-forma en $U \subseteq \mathbb{R}^3$ es, como se dijo anteriormente, una expresión del tipo

$$\omega = \sum_{i,j=1}^{3} f_{ij} dx_i dx_j$$

que, al desarrollarla se ve como

$$\begin{aligned}\omega = {} & f_{11} dx_1 dx_1 + f_{22} dx_2 dx_2 + f_{33} dx_3 dx_3 + f_{12} dx_1 dx_2 \\ & + f_{21} dx_2 dx_1 + f_{13} dx_1 dx_3 + f_{31} dx_3 dx_1 + f_{23} dx_2 dx_3 + f_{32} dx_3 dx_2\end{aligned}$$

Los tres primeros sumandos de esta expresión son entonces iguales a cero. Con $dx_2 dx_1 = -dx_1 dx_2$, $dx_3 dx_1 = -dx_1 dx_3$, $dx_3 dx_2 = -dx_2 dx_3$, la expresión restante queda escrita como

$$\begin{aligned}\omega &= (f_{12} - f_{21}) dx_1 dx_2 + (f_{13} - f_{31}) dx_1 dx_3 + (f_{23} - f_{32}) dx_2 dx_3 \\ &= \sum_{1 \leq i < j \leq 3} (f_{ij} - f_{ji}) dx_i dx_j\end{aligned}$$

Nótese que siendo f_{ij} y f_{ji} funciones diferenciables en $U \subseteq \mathbb{R}^3$, su diferencia $g_{ij} = f_{ij} - f_{ji}$ es una nueva función diferenciable. Entonces una 2-forma en $U \subseteq \mathbb{R}^3$ es, en general, una expresión del tipo

$$\omega = \sum_{1 \leq i < j \leq 3} g_{ij} dx_i dx_j$$

Es claro entonces que una 2-forma ω en $U \subseteq \mathbb{R}^n$ $(n \geq 2)$ tiene una expresión del tipo

$$\omega = \sum_{1 \leq i < j \leq n} g_{ij} dx_i dx_j$$

en esta suma aparecen $\binom{n}{2} = \frac{n!}{2!(n-2)!}$ sumandos (o sea la cantidad de maneras como se pueden escoger dos números distintos en un conjunto de n números). Entonces, la expresión general que determina:

a. una 2-forma en $U \subseteq \mathbb{R}^2$ es $\omega = g dx_1 dx_2$, en donde $g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es una función diferenciable.

b. una 2-forma en $U \subseteq \mathbb{R}^3$ es (como se obtuvo anteriormente)

$$\omega = \sum_{1 \leq i < j \leq 3} g_{ij} dx_i dx_j = g_{12} dx_1 dx_2 + g_{13} dx_1 dx_3 + g_{23} dx_2 dx_3$$

c. una 2-forma en $U \subseteq \mathbb{R}^4$ es

$$\begin{aligned}\omega &= \sum_{1 \leq i < j \leq 4} g_{ij} dx_i dx_j \\ &= g_{12} dx_1 dx_2 + g_{13} dx_1 dx_3 + g_{14} dx_1 dx_4 + g_{23} dx_2 dx_3 + g_{24} dx_2 dx_4 + g_{34} dx_3 dx_4\end{aligned}$$

etc.

Antes de dar la definición general de una p-forma en $U \subseteq \mathbb{R}^n$, se harán varias observaciones sobre algunos de estos objetos que ya han aparecido en el texto. Remitámonos a la sección 3 del capítulo 7, donde se definió la integral de línea $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ de un campo vectorial continuo $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$ a lo largo de un camino $\boldsymbol{\lambda}: [a, b] \to U$ de clase $\mathscr{C}^1$. Una manera de denotar esta integral es

$$\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda} = \int_{\boldsymbol{\lambda}} F_1(x) dx_1 + F_2(x) dx_2 + \cdots + F_n(x) dx_n$$

Con esta notación, obsérvese que el "integrando" del lado derecho de esta expresión no es más que la 1-forma en U

$$\omega = F_1(x) dx_1 + F_2(x) dx_2 + \cdots + F_n(x) dx_n$$

Así pues, la integral de línea $\int_{\boldsymbol{\lambda}} \mathbf{F} \cdot d\boldsymbol{\lambda}$ puede ser escrita como $\int_{\boldsymbol{\lambda}} \omega$, en donde ω es la 1-forma anterior.

Por otra parte, en la (parte final de la) sección 2 del capítulo 9 se escribió la integral de superficie $\iint_{K_f} \mathbf{F}\, dA$, del campo contínuo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$ sobre la superficie K (parametrizada por la función $\mathbf{f}$), como

$$\iint_{K_f} \mathbf{F}\, dA = \iint_{K_f} F_1\, dydz + F_2\, dzdx + F_3\, dxdy$$

De nuevo, observe que el integrando del lado derecho de esta expresión es la 2-forma en U, $\omega = F_1 dydz + F_2 dzdx + F_3 dxdy$, de modo que las integrales de superficie de campos vectoriales son también integrales de formas diferenciales.

Uno de los objetivos que se persigue en este capítulo, como ya se apuntaba en la introducción del mismo, es generalizar los conceptos de integral que han aparecido a lo largo del libro (los cuales a su vez son generalizaciones del concepto de integral de una función real de una variable real estudiado en el primer curso de cálculo). Se pretende, pues, buscar una manera de definir la "integral de algo" que generalice "todas las integrales" que conocemos hasta el momento. Este "algo" serán precisamente las formas diferenciales.

Daremos ahora la definición de una p-forma en el conjunto abierto $U \subseteq \mathbb{R}^n$. Sea $\mathbf{p} \in \mathbb{R}$. Una *p-forma* (pe forma) en $U \subseteq \mathbb{R}^n$ es una expresión del tipo

$$\omega = \sum_I f_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

donde la suma se extiende sobre los elementos del conjunto $I = \{1 \leq i_1, i_2, \ldots, i_p \leq n\}$, y las funciones $f_{i_1 i_2 \ldots i_p} \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ son funciones diferenciables. Si éstas son de clase $\mathscr{C}^k$, se dice que la forma ω es de clase $\mathscr{C}^k$. Nuevamente observamos que si $0 \leq p \leq n$, por la propiedad anticonmutativa del producto de los dx_i, en la suma anterior no aparecerán los sumandos correspondientes a los índices $i_1, i_2, \ldots, i_p$ cuando alguno o algunos de éstos se repitan. Por ejemplo, en una 3-forma en $\mathbb{R}^3$ no aparecerá el sumando $f_{121} dx_1 dx_2 dx_1$ pues $dx_1 dx_2 dx_1 = -dx_1 dx_1 dx_2$ que es igual a cero (estamos pensando en una "asociatividad" en el producto de los dx_i y en que estos símbolos cuando quedan multiplicando por cero se reducen a cero, de modo que $-dx_1 dx_1 dx_2 = -0 dx_2 = 0$). Entonces la expresión de una p-forma en $\mathbb{R}^n$ es

$$\omega = \sum_{1 \leq i_1 < i_2 < \ldots < i_p \leq n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \ldots dx_{i_p}$$

habiendo $\binom{n}{p} = \frac{n!}{p!(n-p)!}$ sumandos en esta expresión, donde $g_{i_1 i_2 \ldots i_p}$ son funciones reales diferenciables definidas en U.

La igualdad entre p-formas en $U \subseteq \mathbb{R}^n$ se define de la manera natural: si ω y ω' son las p-formas

$$\omega = \sum_{1 \leq i_1 < i_2 < \ldots < i_p \leq n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \ldots dx_{i_p}$$

$$\omega' = \sum_{1 \leq i_1 < i_2 < \ldots < i_p \leq n} g'_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \ldots dx_{i_p}$$

entonces ω es igual a ω', lo cual se escribe como $\omega = \omega'$, si

$$g_{i_1 i_2 \ldots i_p} = g'_{i_1 i_2 \ldots i_p}$$

para todo conjunto de índices $i_1, i_2, \ldots, i_p$.

Una p-forma en $U \subseteq \mathbb{R}^n$ para la cual todas las funciones $g_{i_1 i_2 \ldots i_p}$ son iguales a (la función) cero, o bien, todos los $dx_{i_1} dx_{i_2} \ldots dx_{i_p}$ son cero, se llama "p-forma cero" y se escribe simplemente como $\omega = 0$.

Ejemplo 1. Los casos interesantes de p-formas en $U \subseteq \mathbb{R}^n$ ocurren cuando $1 \leq p \leq n$, donde la p-forma está constituida por una suma de $\binom{n}{p}$ términos. Si $p > n$, una p-forma en $U \subseteq \mathbb{R}^n$ es cero. En efecto, según la expresión general, en este caso se tendría una forma ω escrita como

$$\omega = \sum_I f_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

donde la suma se hace sobre todo el conjunto de p enteros $i_1, i_2, \ldots, i_p$ entre 1 y n. Puesto que $p > n$, *cualquier* elección de p enteros entre 1 y n, contiene *al menos una repetición*, en cuyo caso entonces $dx_{i_1} dx_{i_2} \cdots dx_{i_p}$ es igual a cero. De esta manera todos los sumandos de ω son iguales a cero. Esta forma es, por lo tanto, la forma cero. Por ejemplo, una 3-forma en $U \subseteq \mathbb{R}^2$ se vería como

$$\omega = \sum_I f_{i_1 i_2 i_3} dx_{i_1} dx_{i_2} dx_{i_3}$$

en que la suma se efectúa sobre todo el conjunto de 3 enteros escogidos del conjunto $\{1, 2\}$, es decir

$$\begin{aligned} \omega = {} & f_{111} dx_1 dx_1 dx_1 + f_{222} dx_2 dx_2 dx_2 + f_{121} dx_1 dx_2 dx_1 + f_{112} dx_1 dx_1 dx_2 \\ & + f_{221} dx_2 dx_2 dx_1 + f_{212} dx_2 dx_1 dx_2 + f_{211} dx_2 dx_1 dx_1 + f_{122} dx_1 dx_2 dx_2 \end{aligned}$$

Nótese que cada uno de estos sumandos es igual a cero, pues dx_1 ó dx_2 aparece repetida. Entonces $\omega = 0$. ■

Ejemplo 2. Una 1-forma en el conjunto abierto $U \subseteq \mathbb{R}^1 = \mathbb{R}$ (por ejemplo un intervalo abierto) es una expresión del tipo

$$\omega = f dx$$

donde $f: U \subseteq \mathbb{R} \to \mathbb{R}$ es una función diferenciable. Una 2-forma en el conjunto abierto $U \subseteq \mathbb{R}^2$ es una expresión del tipo

$$\omega = g\, dx_1 dx_2$$

donde $g: U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es una función diferenciable. En este caso se tiene un solo sumando correspondiente a la manera de escoger dos números $i < j$ en el conjunto $\{1, 2\}$. De modo semejante, una 3-forma en $U \subseteq \mathbb{R}^3$ es una expresión en la que aparece el único sumando $g dx_1 dx_2 dx_3$, con $g: U \subseteq \mathbb{R}^3 \to \mathbb{R}$ diferenciable. En general, una n-forma (de clase $\mathscr{C}^k$) en $U \subseteq \mathbb{R}^n$ es una expresión del tipo

$$\omega = g\, dx_1 dx_2 \cdots dx_n$$

donde $g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función diferenciable (de clase $\mathscr{C}^k$, respectivamente). ■

Nótese que las 1-formas en $U \subseteq \mathbb{R}$ son los integrandos que aparecen en las integrales del primer curso de cálculo $\int_{[a,b]} f(x)\, dx = \int_{[a,b]} \omega$, donde $[a, b] \subset U$. Las 2-formas en $U \subseteq \mathbb{R}^2$ son los integrandos de las integrales dobles del capítulo 6 y las 3-formas en $U \subseteq \mathbb{R}^3$ son los integrandos de las integrales triples del capítulo 6. En general, las n-formas en $U \subseteq \mathbb{R}^n$ son los integrandos de las integrales n-múltiples, las cuales se efectúan en regiones $\Omega \subset U$.

Ejemplo 3. Una 3-forma en $\mathbb{R}^4$ es una expresión del tipo

$$\begin{aligned}\omega &= \sum_{1\leq i_1<i_2<i_3\leq 4} g_{i_1i_2i_3}dx_{i_1}dx_{i_2}dx_{i_3} \\ &= g_{123}dx_1dx_2dx_3 + g_{124}dx_1dx_2dx_4 + g_{134}dx_1dx_3dx_4 + g_{234}dx_2dx_3dx_4\end{aligned}$$

Por ejemplo, $\omega = (x_1^2 \operatorname{sen} x_2 + x_3)dx_1dx_3dx_4 + x_2x_4dx_2dx_3dx_4$ es una 3-forma en $\mathbb{R}^4$ de clase $\mathscr{C}^\infty$ ■

Pasemos a considerar ahora las operaciones básicas que se efectúan entre formas.

Sean ω, η dos p-formas en el conjunto abierto $U \subseteq \mathbb{R}^n$, digamos

$$\omega = \sum_{1\leq i_1<i_2<\ldots<i_p\leq n} g_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

$$\eta = \sum_{1\leq i_1<i_2<\ldots<i_p\leq n} h_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

Se define la *suma de* ω y η, denotada por $\omega + \eta$, como la p-forma en U dada por

$$\omega + \eta = \sum_{1\leq i_1<i_2<\ldots<i_p\leq n} \left(g_{i_1i_2\ldots i_p} + h_{i_1i_2\ldots i_p}\right)dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

Se define el *producto de la p-forma ω por la* 0-*forma f en U*, denotado por $f\omega$, como la p-forma en U dada por

$$f\omega = \sum_{1\leq i_1<i_2<\ldots<i_p\leq n} fg_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

Ciertamente estas definiciones son muy naturales. Una operación más general que la multiplicación de p-formas por 0-formas dada anteriormente, es la multiplicación de una p-forma por una q-forma: sea ω una p-forma en U y η una q-forma en U, digamos

$$\omega = \sum_{1\leq i_1<i_2<\ldots<i_p\leq n} g_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

$$\eta = \sum_{1\leq j_1<j_2<\ldots<j_q\leq n} h_{j_1j_2\ldots j_q}dx_{j_1}dx_{j_2}\ldots dx_{j_q}$$

Se define el *producto de* ω y η, denotado por $\omega\eta$ como la $(p+q)$-forma en U dada por

$$\omega\eta = \sum_{\substack{1\leq i_1<i_2<\ldots<i_p\leq n\\ 1\leq j_1<j_2<\ldots<j_q\leq n}} g_{i_1i_2\ldots i_p}h_{j_1j_2\ldots j_q}dx_{i_1}dx_{i_2}\cdots dx_{i_p}dx_{j_1}dx_{j_2}\cdots dx_{j_q}$$

(Es claro que si $q = 0$, se tiene la definición anterior para este caso). Obsérvese que "muchos" de los sumandos que aparecen en el producto $\omega\eta$ de la última expresión se pueden anular: basta con que alguno de los índices $i_1, i_2, \ldots, i_p$ coincida con alguno de los índices $j_1, j_2, \ldots, j_q$ para que $dx_{i_1}dx_{i_2}\ldots dx_{i_p}dx_{j_1}dx_{j_2}\ldots dx_{j_q}$ sea igual a cero. Nótese también que este producto se puede obtener de modo formal multiplicando cada uno de los sumandos de ω por cada uno de los sumandos de η

(como un producto ordinario de polinomios), *pero respetando* el orden de las dx's: primero se deben escribir las de ω y luego las de η. De hecho, este producto entre formas *no es conmutativo*. Por ejemplo, considere las 1-formas en $U \subseteq \mathbb{R}^2$, $\omega = g_1dx_1 + g_2dx_2$, $\eta = h_1dx_1 + h_2dx_2$. Se tiene

$$\begin{aligned}\omega\eta &= (g_1dx_1 + g_2dx_2)(h_1dx_1 + h_2dx_2)\\ &= g_1h_1dx_1dx_1 + g_1h_2dx_1dx_2 + g_2h_1dx_2dx_1 + g_2h_2dx_2dx_2\\ &= 0 + g_1h_2dx_1dx_2 + g_2h_1dx_2dx_1 + 0\\ &= (g_1h_2 - g_2h_1)dx_1dx_2\end{aligned}$$

En tanto que

$$\begin{aligned}\eta\omega &= (h_1dx_1 + h_2dx_2)(g_1dx_1 + g_2dx_2)\\ &= h_1g_1dx_1dx_1 + h_1g_2dx_1dx_2 + h_2g_1dx_2dx_1 + h_2g_2dx_2dx_2\\ &= 0 + h_1g_2dx_1dx_2 + h_2g_1dx_2dx_1 + 0\\ &= (h_1g_2 - h_2g_1)dx_1dx_2\end{aligned}$$

de donde se ve que $\omega\eta \neq \eta\omega$. Se tiene, de hecho, $\omega\eta = -\eta\omega$. Esta propiedad se explicará más adelante.

Ejemplo 4. Sea ω la 1-forma en $\mathbb{R}^2$, $\omega = x_1dx_1 - x_1x_2dx_2$, y sea η la 1-forma en $\mathbb{R}^2$, $\eta = 3x_2dx_1 + 5x_1dx_2$. La suma de ω y η es la 1-forma en $\mathbb{R}^2$

$$\omega + \eta = (x_1 + 3x_2)dx_1 + (-x_1x_2 + 5x_1)dx_2$$

El producto de ω por η es la $(1 + 1)$-forma en $\mathbb{R}^2$ que obtenemos como

$$\begin{aligned}\omega\eta &= (x_1dx_1 - x_1x_2dx_2)(3x_2dx_1 + 5x_1dx_2)\\ &= (x_1)(3x_2)dx_1dx_1 + (x_1)(5x_1)dx_1dx_2 + (-x_1x_2)(3x_2)dx_2dx_1 + (-x_1x_2)(5x_1)dx_2dx_2\\ &= 0 + 5x_1^2dx_1dx_2 + (-x_1x_2)(3x_2)dx_2dx_1 + 0\\ &= 5x_1^2dx_1dx_2 - (-x_1x_2)(3x_2)dx_1dx_2 = (5x_1^2 + 3x_1x_2^2)dx_1dx_2\end{aligned}$$

■

Ejemplo 5. Sea ω la 2-forma en $\mathbb{R}^3$

$$\omega = x_1dx_1dx_2 + (x_1 - 3x_2x_3)dx_1dx_3 + x_1^2x_3dx_2dx_3$$

y sea η la 2-forma en $\mathbb{R}^3$

$$\eta = x_2x_3dx_1dx_2 + x_1x_2x_3dx_1dx_3 + (x_1 + x_2 + x_3)dx_2dx_3$$

La suma de estas formas es la 2-forma en $\mathbb{R}^3$

$$\omega + \eta = (x_1 + x_2x_3)dx_1dx_2 + (x_1 - 3x_2x_3 + x_1x_2x_3)dx_1dx_3 + (x_1^2x_3 + x_1 + x_2 + x_3)dx_2dx_3$$

El producto de ω por la 0-forma en $\mathbb{R}^3$, $f(x_1, x_2, x_3) = x_2x_3 \operatorname{sen} x_1$ es

$$f\omega = (x_2x_3 \operatorname{sen} x_1)x_1dx_1dx_2 + (x_2x_3 \operatorname{sen} x_1)(x_1 - 3x_2x_3)dx_1dx_3 + (x_2x_3 \operatorname{sen} x_1)x_1^2x_3dx_2dx_3$$

Consideremos ahora la 1-forma en $\mathbb{R}^3$, τ, dada por

$$\tau = 3x_2dx_1 + 5dx_2 - x_1x_3dx_3$$

El producto de ω por τ será una $(2 + 1)$-forma en $\mathbb{R}^3$ (tendrá entonces un solo sumando —ver ejemplo 2). Podemos proceder formalmente como sigue

$$\begin{aligned}
\omega\tau &= \big(x_1dx_1dx_2 + (x_1 - 3x_2x_3)dx_1dx_3 + x_1^2x_3dx_2dx_3\big)\big(3x_2dx_1 + 5dx_2 - x_1x_3dx_3\big)\\
&= (x_1)(3x_2)dx_1dx_2dx_1 + (x_1)(5)dx_1dx_2dx_2 + (x_1)(-x_1x_3)dx_1dx_2dx_3\\
&+ (x_1 - 3x_2x_3)(3x_2)dx_1dx_3dx_1 + (x_1 - 3x_2x_3)(5)dx_1dx_3dx_2 + (x_1 - 3x_2x_3)(-x_1x_3)dx_1dx_3dx_3\\
&+ (x_1^2x_3)(3x_2)dx_2dx_3dx_1 + (x_1^2x_3)(5)dx_2dx_3dx_2 + (x_1^2x_3)(-x_1x_3)dx_2dx_3dx_3\\
&= (x_1)(-x_1x_3)dx_1dx_2dx_3 + (x_1 - 3x_2x_3)(5)dx_1dx_3dx_2 + (x_1^2x_3)(3x_2)dx_2dx_3dx_1\\
&= (x_1)(-x_1x_3)dx_1dx_2dx_3 - (x_1 - 3x_2x_3)(5)dx_1dx_2dx_3 + (x_1^2x_3)(3x_2)dx_1dx_2dx_3\\
&= \big(-x_1^2x_3 - 5x_1 + 15x_2x_3 + 3x_1^2x_2x_3\big)dx_1dx_2dx_3
\end{aligned}$$ ■

En el siguiente teorema se recogen algunas propiedades importantes de la suma de p-formas y del producto de éstas por 0-formas.

Teorema 10.1.1 Sean ω_1, ω_2 y ω_3 tres p-formas en $U \subseteq \mathbb{R}^n$, y ϕ y ψ dos 0-formas en U. Entonces

a. $\omega_1 + \omega_2 = \omega_2 + \omega_1$

b. $\omega_1 + (\omega_2 + \omega_3) = (\omega_1 + \omega_2) + \omega_3$

c. $\omega_1 + 0 = \omega_1$ (en donde 0 es la p-forma cero)

d. dada la p-forma ω_1, existe una p-forma $(-\omega_1)$ tal que $\omega_1 + (-\omega_1) = 0$

e. $\phi(\omega_1 + \omega_2) = \phi\omega_1 + \phi\omega_2$

f. $(\phi + \psi)\omega_1 = \phi\omega_1 + \psi\omega_1$

g. $\phi(\psi\omega_1) = \psi(\phi\omega_1) = (\phi\psi)\omega_1$

h. $1\omega_1 = \omega_1$ (donde 1 es la 0-forma $1\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, $1(x) = 1\ \forall x \in U$)

Demostración. Se trata de verificaciones de rutina. Sólo se hará la demostración de los incisos **a** y **f**, y el resto queda como ejercicio para el lector. Sean

$$\omega_1 = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} g_{i_1i_2\ldots i_p} dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

$$\omega_2 = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} h_{i_1i_2\ldots i_p} dx_{i_1}dx_{i_2}\ldots dx_{i_p}$$

Entonces

$$\begin{aligned}
\omega_1 + \omega_2 &= \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \big(g_{i_1i_2\ldots i_p} + h_{i_1i_2\ldots i_p}\big)dx_{i_1}dx_{i_2}\ldots dx_{i_p}\\
&= \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \big(h_{i_1i_2\ldots i_p} + g_{i_1i_2\ldots i_p}\big)dx_{i_1}dx_{i_2}\ldots dx_{i_p}\\
&= \omega_2 + \omega_1
\end{aligned}$$

lo que muestra **a**. Por otra parte

$$\begin{aligned}(\phi+\psi)\omega_1 &= \sum_{1\le i_1<i_2<\cdots<i_p\le n}(\phi+\psi)g_{i_1i_2\cdots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}\\ &= \sum_{1\le i_1<i_2<\cdots<i_p\le n}\left(\phi g_{i_1i_2\cdots i_p}+\psi g_{i_1i_2\cdots i_p}\right)dx_{i_1}dx_{i_2}\cdots dx_{i_p}\\ &= \sum_{1\le i_1<i_2<\cdots<i_p\le n}\phi g_{i_1i_2\cdots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}\\ &\quad+\sum_{1\le i_1<i_2<\cdots<i_p\le n}\psi g_{i_1i_2\cdots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}\\ &= \phi\omega_1+\psi\omega_1\end{aligned}$$

lo que muestra **f**. Q.E.D.

Nótese que el comportamiento algebraico del conjunto de las p-formas en $U\subseteq\mathbb{R}^n$, respecto del conjunto de las 0-formas en U, es exactamente el mismo que el comportamiento algebraico de las n-adas ordenadas de $\mathbb{R}^n$, respecto de los números reales. Este último le dio a $\mathbb{R}^n$ la estructura de "espacio vectorial sobre $\mathbb{R}$". En la nota al pie de página en la sección 1 del capítulo 1 se estableció lo que es, en general, un espacio vectorial sobre un campo K (el de los escalares). Por comparación, se ve que todas las propiedades se cumplen considerando los vectores como p-formas y los escalares como 0-formas. La única diferencia es que en nuestro caso las 0-formas *no constituyen un campo* (no todas las 0-formas no nulas poseen inversos multiplicativos). La estructura algebraica que poseen las 0-formas es una estructura más débil que la de campo, llamada *anillo* (conmutativo con unidad). Cuando se tiene un conjunto V y un anillo A ("de escalares"), y se cumplen todas las propiedades 1–8 de la definición de espacio vectorial, se dice que V es un *módulo* (sobre A). En nuestro caso, entonces, el conjunto de las p-formas en $U\subseteq\mathbb{R}^n$, que denotaremos como $\Omega^p(U)$, forma un módulo sobre el anillo de las 0-formas en U, $\Omega^0(U)$.

En el teorema siguiente se establecen algunas propiedades adicionales de las operaciones entre p-formas ya definidas.

Teorema 10.1.2 Sean ω_1 y ω_2 dos p-formas en $U\subseteq\mathbb{R}^n$, η una q-forma en U y τ una r-forma en U. Entonces:

a. $(\omega_1+\omega_2)\eta=\omega_1\eta+\omega_2\eta$

b. $\eta(\omega_1+\omega_2)=\eta\omega_1+\eta\omega_2$

c. $\omega_1(\eta\tau)=(\omega_1\eta)\tau$

d. $\eta\tau=(-1)^{qr}\tau\eta$

Demostración. Los incisos **a**, **b** y **c** son de verificación inmediata, y la dejamos como ejercicio para el lector. Veamos el argumento que prueba **d**. Sean

$$\eta=\sum_{1\le i_1<i_2<\cdots<i_q\le n}g_{i_1i_2\cdots i_q}dx_{i_1}dx_{i_2}\cdots dx_{i_q}$$

$$\tau=\sum_{1\le j_1<j_2<\cdots<j_r\le n}h_{j_1j_2\cdots j_r}dx_{j_1}dx_{j_2}\cdots dx_{j_r}$$

Se tiene

$$\eta\tau = \sum_{\substack{1\leq i_1<i_2<\cdots<i_q\leq n\\ 1\leq j_1<j_2<\cdots<j_r\leq n}} g_{i_1 i_2\ldots i_q} h_{j_1 j_2\ldots j_r} dx_{i_1} dx_{i_2}\ldots dx_{i_q} dx_{j_1} dx_{j_2}\ldots dx_{j_r}$$

Por otra parte

$$\tau\eta = \sum_{\substack{1\leq i_1<i_2<\cdots<i_q\leq n\\ 1\leq j_1<j_2<\cdots<j_r\leq n}} h_{j_1 j_2\ldots j_r} g_{i_1 i_2\ldots i_q} dx_{j_1} dx_{j_2}\ldots dx_{j_r} dx_{i_1} dx_{i_2}\ldots dx_{i_q}$$

Para pasar de la expresión de $\eta\tau$ a la de $\tau\eta$, podemos pensar en un proceso de "ir pasando las dx_j's a través de las dx_i's", tomando en cuenta que cada cambio de posición de 2 de estos símbolos, produce un cambio de signo en el término correspondiente. Por ejemplo

$$\begin{aligned} dx_{i_1} dx_{i_2}\ldots dx_{i_q} dx_{j_1} dx_{j_2}\ldots dx_{j_r} &= -dx_{i_1} dx_{i_2}\ldots dx_{j_1} dx_{i_q} dx_{j_2}\ldots dx_{j_r} \\ &= \ldots = (-1)^{q-1} dx_{i_1} dx_{j_1} dx_{i_2}\ldots dx_{i_q} dx_{j_2}\ldots dx_{j_r} \\ &= (-1)^q dx_{j_1} dx_{i_1} dx_{i_2}\ldots dx_{i_q} dx_{j_2}\ldots dx_{j_r} \end{aligned}$$

Se ve entonces que para llevar dx_{j_1} a la primera posición, éste tuvo que dar q "saltos" (los correspondientes a cada una de las dx_i's), contribuyendo cada uno de estos con un signo menos. Este proceso se repite ahora con dx_{j_2}, teniéndose entonces que

$$dx_{i_1} dx_{i_2}\cdots dx_{i_q} dx_{j_1} dx_{j_2}\cdots dx_{j_r} = (-1)^{2q} dx_{j_1} dx_{j_2} dx_{i_1} dx_{i_2}\cdots dx_{i_q} dx_{j_3}\cdots dx_{j_r}$$

Ahora resulta claro que después de pasar las r dx_j's a través de las q dx_i's, se tendrá una contribución de qr signos menos en la expresión resultante. Es decir, se tiene

$$dx_{i_1} dx_{i_2}\cdots dx_{i_q} dx_{j_1} dx_{j_2}\cdots dx_{j_r} = (-1)^{qr} dx_{j_1} dx_{j_2}\cdots dx_{j_r} dx_{i_1} dx_{i_2}\cdots dx_{i_q}$$

por lo que

$$\begin{aligned} \eta\tau &= \sum_{\substack{1\leq i_1<i_2<\cdots<i_q\leq n\\ 1\leq j_1<j_2<\cdots<j_r\leq n}} g_{i_1 i_2\ldots i_q} h_{j_1 j_2\ldots j_r} dx_{i_1} dx_{i_2}\cdots dx_{i_q} dx_{j_1} dx_{j_2}\cdots dx_{j_r} \\ &= \sum_{\substack{1\leq i_1<i_2<\cdots<i_q\leq n\\ 1\leq j_1<j_2<\cdots<j_r\leq n}} h_{j_1 j_2\ldots j_r} g_{i_1 i_2\ldots i_q} (-1)^{qr} dx_{j_1} dx_{j_2}\cdots dx_{j_r} dx_{i_1} dx_{i_2}\cdots dx_{i_q} \\ &= (-1)^{qr}\tau\eta \end{aligned}$$

como se quería establecer. Q.E.D.

En la próxima sección empezaremos el estudio de una operación muy importante de formas diferenciales, relacionada con el aspecto diferencial de su manejo.

Ejercicios (Capítulo 10, Sección 1)

En los ejercicios 1–10, identifique las expresiones dadas como p-formas (p =?) definidas en el conjunto abierto $U \subseteq \mathbb{R}^n$ (U =?, n =?).

1. $f(x) = 3x^2 + \operatorname{sen} x$

2. $f(x, y, z) = x^2y^3z^4$

3. $\omega = dx_1 + dx_2$

4. $\omega = x_1x_2dx_1dx_2$

5. $\omega = (3x_1 + 2x_2x_3)dx_1dx_2dx_3 + x_2x_4dx_1dx_2dx_4 - x_3x_4dx_1dx_3dx_4$

6. $\omega = x_5x_4x_3x_2x_1dx_1dx_2dx_3dx_4dx_5$

7. $\omega = (\ln x_1)(\ln x_2)dx_1dx_4dx_6 + x_1^4x_2^6dx_2dx_5dx_6$

8. $\omega = \sqrt{x_1^2 + x_2^2 + x_3^2}dx_1dx_3dx_4 + dx_1dx_2dx_3$

9. $\omega = \exp(x_2)dx_1dx_3dx_4dx_7 + x_2x_4x_7dx_4dx_5dx_6dx_7$

10. $\omega = x_1 \operatorname{sen} x_3dx_1dx_4dx_8dx_{12} + x_2x_4x_6x_8dx_2dx_5dx_9dx_{10}$

Para cada pareja de p-formas ω y η dadas en los ejercicios 11–15, obtenga las formas $\omega + \eta$ y $\omega\eta$.

11. $\omega = x_1$, $\eta = \operatorname{sen} x_1$

12. $\omega = x_2dx_1 + x_1dx_2$, $\eta = x_1dx_1 - x_2dx_2$

13. $\omega = dx_1dx_2 + 3x_3dx_2dx_3$, $\eta = x_1x_2x_3dx_1dx_3$

14. $\omega = x_1x_3dx_1dx_3dx_5 - (3x_1 - x_2)dx_1dx_3dx_7$, $\eta = x_2x_4x_7dx_2dx_4dx_6$

15. $\omega = (x_1x_2 + \operatorname{sen}^2 x_3 - \cos^2 x_4)dx_1dx_4dx_5dx_6 + \sqrt{x_1^2 + \ln^4 x_2^2}dx_2dx_3dx_5dx_6$
$+ x_3x_4 \exp(x_1 + x_4)dx_1dx_2dx_3dx_4 + (x_2^2 + x_4^2 - x_6)dx_1dx_3dx_4dx_5$
$\eta = x_1x_2x_3x_5dx_1dx_3dx_5dx_6 + (x_1 + x_2^2 + x_3^3)dx_2dx_4dx_5dx_6$

En los ejercicios 16–20, obtenga los productos $\omega\,\eta$ y $\eta\omega$ de las formas ω y η dadas. Compruebe que se satisfaga la propiedad del inciso **d** del teorema 10.1.2.

16. $\omega = x_1 + x_2$, $\eta = x_1dx_1dx_3 - x_3dx_2dx_3$

17. $\omega = 3x_1x_2dx_1dx_3 - x_4dx_2dx_4$, $\eta = 2x_1dx_1 - 6x_2dx_2 + dx_4$

18. $\omega = 2x_3dx_1dx_4dx_5 - x_2dx_2dx_3dx_4$, $\eta = x_1x_3dx_1dx_5 - x_2dx_2dx_4 + dx_2dx_3$

19. $\omega = dx_1dx_3dx_7 - dx_2dx_3dx_8 + dx_1dx_2dx_3$, $\eta = dx_1dx_4 + dx_2dx_5 - dx_5dx_7$

20. $\omega = 2dx_1dx_5 - 4dx_2dx_4 + x_1dx_3dx_8 - x_7x_8dx_7dx_8 + x_1x_3x_7dx_2dx_7$
$\eta = 2x_1x_5dx_2dx_3 - 4dx_1dx_5 + x_4dx_4dx_7 - x_7x_8dx_7dx_8 + x_2x_4x_8dx_1dx_8$

21. Considere las 1-formas en $\mathbb{R}^3$, $\omega = x_1x_2dx_1 + x_2x_3dx_2 + x_3dx_3$, $\eta = x_1dx_1 + x_2x_3dx_3$, $\rho = x_1x_2x_3dx_1 - dx_2$. Obtenga:

a. $x_2x_3\omega - x_2\eta + x_1\rho$

b. $\omega(\eta + 3\rho)$

c. $(\eta + 3\rho)\omega$

d. $(\omega + 2\eta)(\omega - x_2\rho)$

e. $\omega\eta\rho$

f. $\omega\omega\eta\rho$

g. $\omega\omega\eta\eta\rho\rho$

22. Sean ω y η dos 2-formas en $\mathbb{R}^3$. Obtenga la forma $\omega\eta$.

23. Sea A la matriz $\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$, donde f_{ij} son funciones reales diferenciables definidas en $\mathbb{R}^2$. Considere las 1-formas en $\mathbb{R}^2$

$$\omega_1 = f_{11}dx_1 + f_{21}dx_2, \qquad \omega_2 = f_{12}dx_1 + f_{22}dx_2$$

Demuestre que $\omega_1\omega_2 = \det A\, dx_1dx_2$.

(*) **24.** (Generalización del ejercicio anterior). Sea A la matriz $n \times n$, $A = (f_{ij})_{i,j=1,2,\ldots,n}$, donde f_{ij} son funciones reales diferenciables definidas en $\mathbb{R}^n$. Considere las 1-formas en $\mathbb{R}^n$, $\omega_j = \sum_{i=1}^{n} f_{ij}dx_i$, $j = 1, 2, \ldots, n$. Demuestre que $\omega_1\omega_2\cdots\omega_n = \det A\, dx_1dx_2\cdots dx_n$.

10.2 La diferencial exterior

En esta sección estudiaremos una operación que se efectúa en las formas, llamada diferencial exterior, que será cómo "derivar" formas, para producir nuevas formas. Así como en la sección anterior se identificaron algunas formas que ya habían aparecido en otras partes del libro, veremos aquí también que algunas diferenciales exteriores de formas ya son conocidas. Las formas diferenciales con que trabajaremos en esta sección serán de clase $\mathscr{C}^2$.

Para comenzar, empecemos por considerar una 0-forma ω en $U \subseteq \mathbb{R}^n$. Esta es entonces una función (de clase $\mathscr{C}^2$) $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. En la sección 11 del capítulo 2 se definió la diferencial de f, denotada por df, como

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n$$

donde las derivadas parciales están evaluadas en $x \in U$. Los símbolos dx's que aparecen en esta expresión fueron llamados diferenciales de x. Viendo a estos como los símbolos, escritos de la misma manera, que se presentaron en la sección anterior, se ve que df se puede identificar con una 1-forma en U. Este es, de hecho, el primer ejemplo de los que definiremos más adelante como diferencial exterior de la 0-forma ω. Ya en él se ve que la diferencial de una 0-forma es una 1-forma. Acontecerá en general que esta operación actúa sobre p-formas para producir $(p+1)$-formas.

Definimos entonces formalmente la diferencial exterior de una 0-forma ω en U, dada por la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$, denotada por $d\omega$, como la 1-forma en U

$$d\omega = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n$$

Por ejemplo, la diferencial exterior de la 0-forma en $\mathbb{R}^3 \omega = x_1^2 + 2x_2^3x_3$ es la 1-forma en $\mathbb{R}^3$

$$d\omega = 2x_1dx_1 + 6x_2^2x_3dx_2 + 2x_2^3dx_3$$

Desde esta perspectiva, es posible "provocar" un significado al símbolo dx_i, viendo a éste (presentado en la sección anterior como una 1-forma) como la diferencial exterior de una cierta 0-forma. En efecto, considere la "función proyección i-ésima", $x_i: \mathbb{R}^n \to \mathbb{R}$, $x_i(y_1, y_2, \ldots, y_n) = y_i$ = i-ésima coordenada de $(y_1, y_2, \ldots, y_n)$. Ciertamente ésta es una 0-forma en $\mathbb{R}^n$, cuya diferencial exterior es, según lo definido previamente

$$d(x_i) = \frac{\partial x_i}{\partial y_1}dx_1 + \frac{\partial x_i}{\partial y_2}dx_2 + \cdots + \frac{\partial x_i}{\partial y_n}dx_n$$

Es claro que $\frac{\partial x_i}{\partial y_j}$ será igual a 1 cuando $j = i$, siendo cero para $j \neq i$. Se obtiene entonces que

$$d(x_i) = dx_i$$

Y así, el símbolo dx_i (o bien, la 1-forma dx_i) es la diferencial exterior de la 0-forma x_i.

Consideremos el caso general. Sea ω la p-forma en $U \subseteq \mathbb{R}^n$

$$\omega = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

donde, como sabemos, las funciones $g_{i_1 i_2 \ldots i_p}$ son funciones de clase $\mathscr{C}^2$ definidas en el abierto U de $\mathbb{R}^n$. Viendo a éstas como 0-formas en U, definimos la *diferencial exterior* de la p-forma ω, denotada por $d\omega$, como la $(p+1)$-forma en U

$$d\omega = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} dg_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

o bien, en forma más explícita

$$d\omega = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j}(g_{i_1 i_2 \ldots i_p})dx_j \right) dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

Nótese que "muchos" de estos $n\binom{n}{p}$ términos que aparecen en $d\omega$ se pueden hacer cero: basta que el índice j de dx_j esté contenido dentro de los p enteros $i_1, i_2, \ldots, i_p$ para que $dx_j dx_{i_1} dx_{i_2} \cdots dx_{i_p}$ sea igual a cero.

Ejemplo 1. Sea ω la 1-forma en $\mathbb{R}^3$

$$\omega = (x_1 + 2x_2)dx_1 + 3x_1x_2dx_2 + x_3^2dx_3$$

La diferencial exterior de ω será la 2-forma en $\mathbb{R}^3$

$$\begin{aligned}
d\omega &= dg_1dx_1 + dg_2dx_2 + dg_3dx_3 \\
&= \left(\frac{\partial}{\partial x_1}(x_1 + 2x_2)dx_1 + \frac{\partial}{\partial x_2}(x_1 + 2x_2)dx_2 + \frac{\partial}{\partial x_3}(x_1 + 2x_2)dx_3\right)dx_1 \\
&\quad + \left(\frac{\partial}{\partial x_1}(3x_1x_2)dx_1 + \frac{\partial}{\partial x_2}(3x_1x_2)dx_2 + \frac{\partial}{\partial x_3}(3x_1x_2)dx_3\right)dx_2 \\
&\quad + \left(\frac{\partial}{\partial x_1}(x_3^2)dx_1 + \frac{\partial}{\partial x_2}(x_3^2)dx_2 + \frac{\partial}{\partial x_3}(x_3^2)dx_3\right)dx_3 \\
&= \big(dx_1 + 2dx_2 + 0\big)dx_1 + \big(3x_2dx_1 + 3x_1dx_2 + 0\big)dx_2 + \big(0 + 0 + 2x_3dx_3\big)dx_3 \\
&= dx_1dx_1 + 2dx_2dx_1 + 3x_2dx_1dx_2 + 3x_1dx_2dx_2 + 2x_3dx_3dx_3 \\
&= 0 - 2dx_1dx_2 + 3x_2dx_1dx_2 + 0 + 0 = (3x_2 - 2)dx_1dx_2
\end{aligned}$$

■

Ejemplo 2. Si ω es una 1-forma en $U \subseteq \mathbb{R}$, su diferencial exterior será una 2-forma en $\mathbb{R}$, la cual será entonces la forma cero (ver ejemplo 1 de la sección anterior). En efecto, si $\omega = fdx$, se tiene $d\omega = dfdx = f'dxdx = 0$ Similarmente, si ω es una p-forma en $U \subseteq \mathbb{R}^p$, $d\omega$ será la $(p+1)$-forma cero. ■

Con el siguiente teorema comenzamos a presentar las propiedades que tiene la diferencial exterior.

Teorema 10.2.1 Sean ω_1 y ω_2 dos p-formas en $U \subseteq \mathbb{R}^n$, entonces

$$d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$$

Demostración. Si

$$\omega_1 = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} g_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$

$$\omega_2 = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} h_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$

Entonces, directamente de la definición de diferencial exterior (y de suma de formas), se tiene

$$\begin{aligned}
d(\omega_1 + \omega_2) &= \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} d\big(g_{i_1i_2\ldots i_p} + h_{i_1i_2\ldots i_p}\big)dx_{i_1}dx_{i_2}\cdots dx_{i_p} \\
&= \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\big(g_{i_1i_2\ldots i_p} + h_{i_1i_2\ldots i_p}\big)dx_j\right)dx_{i_1}dx_{i_2}\cdots dx_{i_p} \\
&= \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \left(\sum_{j=1}^{n} \left(\frac{\partial}{\partial x_j}g_{i_1i_2\ldots i_p} + \frac{\partial}{\partial x_j}h_{i_1i_2\ldots i_p}\right)dx_j\right)dx_{i_1}dx_{i_2}\cdots dx_{i_p} \\
&= \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\big(g_{i_1i_2\ldots i_p}\big)dx_j + \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\big(h_{i_1i_2\ldots i_p}\big)dx_j\right)dx_{i_1}dx_{i_2}\cdots dx_{i_p}
\end{aligned}$$

$$= \sum_{1\le i_1<i_2<\cdots<i_p\le n} \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j}(g_{i_1 i_2\cdots i_p})dx_j\right) dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$

$$+ \sum_{1\le i_1<i_2<\cdots<i_p\le n} \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j}(h_{i_1 i_2\cdots i_p})dx_j\right) dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$

$$= d\omega_1 + d\omega_2$$

Q.E.D.

En general, si $\omega_1, \omega_2, \ldots, \omega_k$ son k p-formas en $U \subseteq \mathbb{R}^n$, se tiene

$$d(\omega_1 + \omega_2 + \cdots + \omega_k) = d\omega_1 + d\omega_2 + \cdots + d\omega_k$$

lo cual se prueba fácilmente (por inducción sobre k) aplicando el teorema anterior.

En el teorema siguiente se establece el comportamiento de la diferencial exterior ante un producto de formas.

Teorema 10.2.2 Sea ω una p-forma en $U \subseteq \mathbb{R}^n$ y η una q-forma en U. Entonces

$$d(\omega\eta) = (d\omega)\eta + (-1)^p \omega d\eta$$

Demostración. Nótese que la fórmula que establece el teorema es una igualdad entre la $(p+q+1)$-forma del lado izquierdo, y la suma de la $(p+q+1)$-formas del lado derecho. Consideremos primeramente el caso en que ω y η son 0-formas en U, digamos $h, g\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$. Se tiene

$$d(hg) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(hg)dx_i = \sum_{i=1}^{n}\left(g\frac{\partial h}{\partial x_i} + h\frac{\partial g}{\partial x_i}\right)dx_i$$

$$= \sum_{i=1}^{n} g\frac{\partial h}{\partial x_i}dx_i + \sum_{i=1}^{n} h\frac{\partial g}{\partial x_i}dx_i = g\sum_{i=1}^{n}\frac{\partial h}{\partial x_i}dx_i + h\sum_{i=1}^{n}\frac{\partial g}{\partial x_i}dx_i$$

$$= gdh + hdg$$

Sean $1 \le i_1 < i_2 < \cdots < i_p \le n$ un conjunto dado de p enteros, y $1 \le j_1 < j_2 < \cdots < j_q \le n$ un conjunto dado de q enteros. Consideremos la p-forma ω^* y la q-forma η^* dadas por

$$\omega^* = gdx_{i_1}dx_{i_2}\cdots dx_{i_p}$$
$$\eta^* = hdx_{j_1}dx_{j_2}\cdots dx_{j_q}$$

donde $g, h\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ son funciones de clase $\mathscr{C}^2$ en U, o bien, 0-formas en U. Se tiene

$$d(\omega^*\eta^*) = d\left(ghdx_{i_1}dx_{i_2}\cdots dx_{i_p}dx_{j_1}dx_{j_2}\cdots dx_{j_q}\right) \quad \text{(Por la definición de producto de formas)}$$

$$= d(gh)dx_{i_1}dx_{i_2}\cdots dx_{i_p}dx_{j_1}dx_{j_2}\cdots dx_{j_q} \quad \text{(Por la definición de diferencial exterior)}$$

$$= (gdh + hdg)dx_{i_1}dx_{i_2}\cdots dx_{i_p}dx_{j_1}dx_{j_2}\cdots dx_{j_q} \quad \text{(Por el caso previamente demostrado)}$$

$$= dgdx_{i_1}dx_{i_2}\cdots dx_{i_p}hdx_{j_1}dx_{j_2}\cdots dx_{j_q}$$

$$+ dh(gdx_{i_1}dx_{i_2}\cdots dx_{i_p})dx_{j_1}dx_{j_2}\cdots dx_{j_q} \quad \text{(Simple reacomodo)}$$

$$= (dg dx_{i_1} dx_{i_2} \cdots dx_{i_p})(h dx_{j_1} dx_{j_2} \cdots dx_{j_q})$$

$$+ (-1)^p (g dx_{i_1} dx_{i_2} \cdots dx_{i_p})(dh dx_{j_1} dx_{j_2} \cdots dx_{j_q})$$

(Pues siendo dh y cada una de las dx_i's 1-formas (p en total de estas últimas), la forma dh se intercambió con cada una de las dx_i's, produciendo un signo menos en cada intercambio. Ver teorema 10.1.2 inciso **d**)

$$= (d\omega^*)\eta^* + (-1)^p \omega^* d\eta^*$$

Consideremos el caso general: sea ω una p-forma en $U \subseteq \mathbb{R}^n$ y η una q-forma en U, digamos

$$\omega = \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

$$\eta = \sum_{1 \le j_1 < j_2 < \cdots < j_q \le n} h_{j_1 j_2 \ldots j_q} dx_{j_1} dx_{j_2} \cdots dx_{j_q}$$

Viendo a estas formas como sumas de formas del tipo ω^* y η^* consideradas anteriormente, digamos

$$\omega = \sum_i \omega_i^*, \qquad \eta = \sum_j \eta_j^*$$

(donde la primera suma va de $i = 1$ hasta $\binom{n}{p}$ y la segunda va de $j = 1$ hasta $\binom{n}{q}$). Se tiene, con el resultado obtenido para las formas ω^* y η^*, junto con la generalización del teorema 10.2.1 comentada al final de éste

$$d(\omega\eta) = d\left[\left(\sum_i \omega_i^*\right)\left(\sum_j \eta_j^*\right)\right] = d\left(\sum_{i,j} \omega_i^* \eta_j^*\right)$$

$$= \sum_{i,j} d(\omega_i^* \eta_j^*) = \sum_{i,j} \left((d\omega_i^*)\eta_j^* + (-1)^p \omega_i^* d\eta_j^*\right)$$

$$= \sum_{i,j} (d\omega_i^*)\eta_j^* + (-1)^p \sum_{i,j} \omega_i^* d\eta_j^*$$

$$= \left(d\left(\sum_i \omega_i^*\right)\right)\left(\sum_j \eta_j^*\right) + (-1)^p \left(\sum_i \omega_i^*\right) d\left(\sum_j \eta_j^*\right)$$

$$= (d\omega)\eta + (-1)^p \omega d\eta$$

Q.E.D.

Ejemplo 3. Consideremos las 1-formas en $\mathbb{R}^2$

$$\omega = x_1 dx_1 - x_1 x_2 dx_2, \eta = 3x_2 dx_1 + 5x_1 dx_2$$

En el ejemplo 4 de la sección 1 se obtuvo que $\omega\eta$ es la 2-forma en $\mathbb{R}^2$

$$\omega\eta = (5x_1^2 + 3x_1 x_2^2) dx_1 dx_2$$

Siendo $\omega\eta$ una 2-forma en $\mathbb{R}^2$, su diferencial exterior $d(\omega\eta)$ será igual a cero (ver ejemplo 2). De modo semejante, siendo ω una 1-forma en $\mathbb{R}^2$, su diferencial exterior $d\omega$ será una 2-forma en $\mathbb{R}^2$,

de modo que el producto de esta 2-forma en $\mathbb{R}^2$ por la 1-forma en $\mathbb{R}^2 \eta$, será una 3-forma en $\mathbb{R}^2$, y por lo tanto, será igual a cero. También el producto $\omega d\eta$, siendo el producto de la 1-forma ω por la 2-forma $d\eta$, será una 3-forma en $\mathbb{R}^2$, y por lo tanto, igual a cero. Se verifica entonces la fórmula del teorema 10.2.2 en el caso trivial, en que ambos miembros son iguales a cero. ■

Ejemplo 4. Considere la 1-forma ω en $\mathbb{R}^4$

$$\omega = x_3x_4dx_1 + x_1x_2x_3dx_3 + x_4dx_4$$

y la 2-forma η en $\mathbb{R}^4$

$$\eta = 2x_1x_4dx_1dx_3 + 5x_2x_3x_4dx_2dx_4$$

El producto $\omega\eta$ es la 3-forma en $\mathbb{R}^4$

$$\begin{aligned}\omega\eta &= (x_3x_4)(5x_2x_3x_4)dx_1dx_2dx_4 + (x_1x_2x_3)(5x_2x_3x_4)dx_3dx_2dx_4 + (x_4)(2x_1x_4)dx_4dx_1dx_3\\ &= 5x_2x_3^2x_4^2dx_1dx_2dx_4 - 5x_1x_2^2x_3^2x_4dx_2dx_3dx_4 + 2x_1x_4^2dx_1dx_3dx_4\end{aligned}$$

Las diferenciales exteriores de estas formas son

$$\begin{aligned}d\omega &= d(x_3x_4)dx_1 + d(x_1x_2x_3)dx_3 + d(x_4)dx_4\\ &= (x_4dx_3 + x_3dx_4)dx_1 + (x_2x_3dx_1 + x_1x_3dx_2 + x_1x_2dx_3)dx_3 + dx_4dx_4\\ &= x_4dx_3dx_1 + x_3dx_4dx_1 + x_2x_3dx_1dx_3 + x_1x_3dx_2dx_3\\ &= (x_2x_3 - x_4)dx_1dx_3 + x_1x_3dx_2dx_3 - x_3dx_1dx_4\\ d\eta &= d(2x_1x_4)dx_1dx_3 + d(5x_2x_3x_4)dx_2dx_4\\ &= (2x_4dx_1 + 2x_1dx_4)dx_1dx_3 + (5x_3x_4dx_2 + 5x_2x_4dx_3 + 5x_2x_3dx_4)dx_2dx_4\\ &= 2x_1dx_4dx_1dx_3 + 5x_2x_4dx_3dx_2dx_4 = 2x_1dx_1dx_3dx_4 - 5x_2x_4dx_2dx_3dx_4\end{aligned}$$

y por último la diferencial exterior de $\omega\eta$ es

$$\begin{aligned}d(\omega\eta) &= d(5x_2x_3^2x_4^2)dx_1dx_2dx_4 - d(5x_1x_2^2x_3^2x_4)dx_2dx_3dx_4 + d(2x_1x_4^2)dx_1dx_3dx_4\\ &= (5x_3^2x_4^2dx_2 + 10x_2x_3x_4^2dx_3 + 10x_2x_3^2x_4dx_4)dx_1dx_2dx_4\\ &\quad - (5x_2^2x_3^2x_4dx_1 + 10x_1x_2x_3^2x_4dx_2 + 10x_1x_2^2x_3x_4dx_3 + 5x_1x_2^2x_3^2dx_4)dx_2dx_3dx_4\\ &\quad + (2x_4^2dx_1 + 4x_1x_4dx_4)dx_1dx_3dx_4\\ &= 10x_2x_3x_4^2dx_3dx_1dx_2dx_4 - 5x_2^2x_3^2x_4dx_1dx_2dx_3dx_4\\ &= (10x_2x_3x_4^2 - 5x_2^2x_3^2x_4)dx_1dx_2dx_3dx_4\end{aligned}$$

Según la fórmula del teorema 10.2.2, esta forma debe ser igual a $(d\omega)\eta + (-1)^1 \omega d\eta$. Hagamos esta cuenta

$$\begin{aligned}(d\omega)\eta - \omega d\eta &= \big((x_2x_3 - x_4)dx_1dx_3 + x_1x_3dx_2dx_3 - x_3dx_1dx_4\big)\big(2x_1x_4dx_1dx_3 + 5x_2x_3x_4dx_2dx_4\big)\\ &\quad - \big(x_3x_4dx_1 + x_1x_2x_3dx_3 + x_4dx_4\big)\big(2x_1dx_1dx_3dx_4 - 5x_2x_4dx_2dx_3dx_4\big)\\ &= (x_2x_3 - x_4)(5x_2x_3x_4)dx_1dx_3dx_2dx_4 + (x_3x_4)(5x_2x_4)dx_1dx_2dx_3dx_4\\ &= \big(-(x_2x_3 - x_4)(5x_2x_3x_4) + (x_3x_4)(5x_2x_4)\big)dx_1dx_2dx_3dx_4\\ &= (10x_2x_3x_4^2 - 5x_2^2x_3^2x_4)dx_1dx_2dx_3dx_4\end{aligned}$$

lo que comprueba la validez de la fórmula del teorema 10.2.2. ■

Tomemos ahora tres formas en $U \subseteq \mathbb{R}^n$, digamos ω una p-forma, η_1 una q_1-forma y η_2 una q_2-forma. Con la propiedad asociativa del producto de formas (teorema 10.1.2c), y la fórmula establecida en el teorema 10.2.2, se tiene que

$$\begin{aligned}d(\omega\eta_1\eta_2) &= d((\omega\eta_1)\eta_2)) = d(\omega\eta_1)\eta_2 + (-1)^{p+q}1\omega\eta_1 d\eta_2\\ &= [(d\omega)\eta_1 + (-1)^p\omega d\eta_1]\eta_2 + (-1)^{p+q}1\omega\eta_1 d\eta_2\\ &= (d\omega)\eta_1\eta_2 + (-1)^p\omega d\eta_1\eta_2 + (-1)^{p+q}1\omega\eta_1 d\eta_2\end{aligned}$$

En general, si ω es una p-forma en $U \subseteq \mathbb{R}^n$, y η_i son q_i-formas en U, $i = 1, 2, \ldots, k$, entonces la diferencial exterior de el producto de estas $(k + 1)$ formas está dada por la siguiente fórmula, que generaliza la establecida en el teorema 10.2.2, cuya demostración dejamos como ejercicio para el lector

$$d(\omega\eta_1\eta_2\ldots\eta_k) = (d\omega)\eta_1\eta_2\ldots\eta_k + \sum_{j=1}^{k}(-1)^{p+q_1+\cdots+q_{j-1}}\omega\eta_1\ldots\eta_{j-1}(d\eta_j)\eta_{j+1}\ldots\eta_k$$

Antes de enunciar el próximo teorema, en el que se establece una nueva propiedad de la diferencial exterior, veamos el ejemplo siguiente.

Ejemplo 5. Considere la 1-forma ω en $\mathbb{R}^4$ del ejemplo anterior

$$\omega = x_3x_4dx_1 + x_1x_2x_3dx_3 + x_4dx_4$$

Su diferencial exterior es

$$d\omega = (x_2x_3 - x_4)dx_1dx_3 + x_1x_3dx_2dx_3 - x_3dx_1dx_4$$

la cual es una 2-forma en $\mathbb{R}^4$. Obtengamos su diferencial exterior. Esta será la 3-forma en $\mathbb{R}^4$

$$\begin{aligned}d(d\omega) &= d(x_2x_3 - x_4)dx_1dx_3 + d(x_1x_3)dx_2dx_3 - d(x_3)dx_1dx_4\\ &= (x_3dx_2 + x_2dx_3 - dx_4)dx_1dx_3 + (x_3dx_1 + x_1dx_3)dx_2dx_3 - dx_3dx_1dx_4\\ &= x_3dx_2dx_1dx_3 - dx_4dx_1dx_3 + x_3dx_1dx_2dx_3 - dx_3dx_1dx_4\\ &= -x_3dx_1dx_2dx_3 - dx_1dx_3dx_4 + x_3dx_1dx_2dx_3 + dx_1dx_3dx_4 = 0\end{aligned}$$

Del mismo modo se tiene para la 2-forma η del ejemplo anterior

$$\begin{aligned} d(d\eta) &= d\left(2x_1dx_1dx_3dx_4 - 5x_2x_4dx_2dx_3dx_4\right) \\ &= d(2x_1)dx_1dx_3dx_4 - d(5x_2x_4)dx_2dx_3dx_4 \\ &= 2dx_1dx_1dx_3dx_4 - (5x_4dx_2 + 5x_2dx_4)dx_2dx_3dx_4 = 0 \end{aligned}$$ ■

La situación presentada en el ejemplo anterior con las formas ω y η es un hecho general que se enuncia y demuestra a continuación.

Teorema 10.2.3 Sea ω una p-forma en $U \subseteq \mathbb{R}^n$. Entonces $d(d\omega) = 0$.

Demostración. Es suficiente mostrar que la fórmula vale para 0-formas en U, pues en tal caso, siendo la forma ω

$$\omega = \sum_{1\leq i_1<i_2<\cdots<i_p\leq n} g_{i_1i_2\cdots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$

se tendría, según la definición de diferencial exterior (aplicada a la forma $d\omega$)

$$d(d\omega) = \sum_{1\leq i_1<i_2<\cdots<i_p\leq n} d(dg_{i_1i_2\cdots i_p})dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$

Siendo las funciones g's 0-formas en U, basta probar que la diferencial exterior de la diferencial exterior de éstas es igual a cero para concluir que $d(d\omega)$ es igual a 0.

Sea entonces $g: U \subseteq \mathbb{R}^n \to \mathbb{R}$ una 0-forma en U. Se tiene

$$\begin{aligned} d(dg) &= d\left(\frac{\partial g}{\partial x_1}dx_1 + \frac{\partial g}{\partial x_2}dx_2 + \cdots + \frac{\partial g}{\partial x_n}dx_n\right) \\ &= d\sum_{i=1}^{n}\frac{\partial g}{\partial x_i}dx_i = \sum_{i=1}^{n} d\left(\frac{\partial g}{\partial x_i}\right)dx_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(\frac{\partial g}{\partial x_i}\right)dx_j\right)dx_i \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 g}{\partial x_j\partial x_i}dx_jdx_i \end{aligned}$$

La suma anterior que contiene en principio n^2 términos se puede considerar como la suma de: **a.** los términos en que $i = j$; **b.** los términos en que $i < j$; **c.** los términos en que $i > j$. Es claro que los términos en que $i = j$ será igual a cero, pues aparecerá dx_idx_i en cada unos de ellos. Separemos entonces la suma anterior como

$$\sum_{i<j}\frac{\partial^2 g}{\partial x_j\partial x_i}dx_jdx_i + \sum_{i>j}\frac{\partial^2 g}{\partial x_j\partial x_i}dx_jdx_i$$

Expresión que es igual a (cambiándole los nombres a los índices en la segunda sumatoria)

$$\sum_{i<j}\frac{\partial^2 g}{\partial x_j\partial x_i}dx_jdx_i + \sum_{i<j}\frac{\partial^2 g}{\partial x_i\partial x_j}dx_idx_j$$

Usando que $dx_j dx_i = -dx_i dx_j$ se tiene finalmente que

$$d(d\omega) = \sum_{i<j} \left(\frac{\partial^2 g}{\partial x_j \partial x_i} - \frac{\partial^2 g}{\partial x_i \partial x_j}\right) dx_j dx_i$$

Recuerde que se está asumiendo clase $\mathscr{C}^2$ para las formas diferenciales en esta sección, por lo que, según el teorema de Schwarz (sobre la igualdad de las derivadas parciales mixtas), todos los términos de la suma anterior son iguales a cero. Q.E.D.

La fórmula del teorema anterior se suele escribir como

$$d^2 = 0$$

Esta es, sin duda, una fórmula de expresión muy simple pero que involucra conceptos y propiedades muy importantes de las formas diferenciales.

Así, la diferencial exterior de la diferencial exterior de cualquier forma es igual a la forma cero. Esta situación nos plantea de manera natural la siguiente pregunta: si la diferencial exterior de alguna p-forma es igual a cero, ¿es esta forma la diferencial exterior de alguna otra $(p-1)$-forma? Es decir, si la p-forma η es tal que $d\eta$ sea igual a cero, ¿es η la diferencial exterior de alguna $(p-1)$-forma ω? Obsérvese que en tal caso se tendría $0 = d\eta = d(d\omega)$, cosa que sabemos que ocurre con todas las formas ω. Esta es una pregunta de respuesta no trivial. Gran parte del capítulo 7 es una prueba de este hecho, pues en él se planteó y se contestó esta pregunta para cuando η es una 1-forma en $\mathbb{R}^n$ (revise de nuevo la historia de la sección 3 del capítulo 7 y vea las preguntas y respuestas ahí planteadas desde el punto de vista de la teoría de formas desarrollada en estas dos secciones).

Para terminar esta sección, estudiaremos algunas identificaciones importantes entre los conjuntos $\Omega^p(U)$ de las p-formas definidas en $U \subseteq \mathbb{R}^n$, con las cuales entenderemos mejor cómo esta teoría de formas difernciales generaliza el trabajo desarrollado en los capítulos anteriores.

Consideremos el conjunto $\Omega^0(U)$ de las 0-formas definidas en $U \subseteq \mathbb{R}^n$. Los elementos de este conjunto son, entonces, funciones diferenciables $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. Cada una de ellas puede ser identificada con una n-forma en U, a saber, la forma $\omega = f dx_1 dx_2 \cdots dx_n$. Así se establece una correspondencia entre los elementos del conjunto $\Omega^0(U)$ y los elementos del conjunto $\Omega^n(U)$, que denotaremos como $T_{0,n}: \Omega^0(U) \to \Omega^n(U)$, dada por

$$T_{0,n}(f) = f dx_1 dx_2 \cdots dx_n$$

Esta función es una biyección, la cual, además, satisface

$$T_{0,n}(f+g) = T_{0,n}(f) + T_{0,n}(g)$$
$$T_{0,n}(\phi f) = \phi T_{0,n}(f)$$

donde $\phi \in \Omega^0(U)$. Dejamos al lector que verifique estos hechos. En otras palabras, $T_{0,n}$ es una transformación lineal biyectiva. Se dice entonces, que $T_{0,n}$ es un *isomorfismo* entre los módulos $\Omega^0(U)$ y $\Omega^n(U)$. Con este isomorfismo podemos identificar los elementos de estos conjuntos.

Consideraremos también la identificación entre los elementos de los conjuntos $\Omega^1(U)$ y $\Omega^{n-1}(U)$, según la función $T_{1,n-1}: \Omega^1(U) \to \Omega^{n-1}(U)$, que asocia a la 1-forma $\omega = f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n$

en $\Omega^1(U)$, la $(n-1)$-forma

$$T_{1,n-1}(\omega) = \sum_{i=1}^{n} (-1)^{i+1} f_i dx_1 dx_2 \cdots dx_{i-1} dx_{i+1} \cdots dx_n$$

Se verifica también que $T_{1,n-1}$ es un isomorfismo estre los módulos $\Omega^1(U)$ y $\Omega^{n-1}(U)$.

Ejemplo 6. Sea ω la 1-forma en $\mathbb{R}^3$, $\omega = x_1dx_1 + x_1x_2x_3dx_2 - x_1x_3dx_3$. Esta forma queda identificada, según el isomorfismo $T_{1,n-1}$, con la 2-forma en $\mathbb{R}^3$

$$\begin{aligned} T_{1,n-1}(\omega) &= (-1)^{1+1}x_1dx_2dx_3 + (-1)^{2+1}x_1x_2x_3dx_1dx_3 - (-1)^{3+1}x_1x_3dx_1dx_2 \\ &= x_1dx_2dx_3 - x_1x_2x_3dx_1dx_3 - x_1x_3dx_1dx_2 \end{aligned}$$ ■

Designemos por $\Phi(U)$ al conjunto de todos los campos vectoriales diferenciables $F\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^n$. En este conjunto definimos las operaciones de suma y producto por funciones diferenciables $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ de la siguiente manera: Si $F = (F_1, F_2, \ldots, F_n)$ y $G = (G_1, G_2, \ldots, G_n)$ son elementos de $\Phi(U)$ y f es un elemento de $\Omega^0(U)$, se tiene

$$\begin{aligned} F + G &= (F_1 + G_1, F_2 + G_2, \ldots, F_n + G_n) \\ fF &= (fF_1, fF_2, \ldots, fF_n) \end{aligned}$$

Se comprueba fácilmente que éste es un módulo, con los escalares como funciones diferenciables (las 0-formas) f.

Podemos identificar los campos vectoriales de $\Phi(U)$ con las 1-formas de $\Omega^1(U)$, de la manera natural: al campo $F = (F_1, F_2, \ldots, F_n) \in \Phi(U)$ le asociamos la 1-forma $\omega = F_1dx_1 + F_2dx_2 + \cdots + F_ndx_n$. Quedan con esto identificados los conjuntos $\Phi(U)$, $\Omega^1(U)$ y $\Omega^{n-1}(U)$. Con esta identificación, por ejemplo, el campo $F(x_1, x_2) = (2x_1 + x_2, x_1 - x_2^2)$ en $\Phi(\mathbb{R}^2)$ queda identificado con la 1-forma en $\mathbb{R}^2$, $\omega = (2x_1 + x_2)dx_1 + (x_1 - x_2^2)dx_2$ y con la 1-forma en $\mathbb{R}^2$, $\eta = (2x_1 + x_2)dx_2 - (x_1 - x_2^2)dx_1$.

Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathscr{C}^2$. El campo

$$\mathbf{F} = \operatorname{grad} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

es un elemento de $\Phi(U)$ que queda identificado con la 1-forma en U

$$\omega = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n$$

Nótese que esta 1-forma no es más que la diferencial exterior de la 0-forma f. Así pues, *el gradiente de la función f se identifica con la 1-forma $\omega = df$.*

Ejemplo 7. Consideremos la 0-forma en $\mathbb{R}^2$, $f(x_1, x_2) = x_1 \cos x_2$. El gradiente de esta función es el campo vectorial en $\mathbb{R}^2$

$$\operatorname{grad} f = (\cos x_2, -x_1 \operatorname{sen} x_2)$$

el cual se identifica con la 1-forma en $\mathbb{R}^2$

$$\omega = \cos x_2 dx_1 - x_1 \operatorname{sen} x_2 dx_2$$

que resulta ser la diferencial exterior de la 0-forma f. ∎

Consideremos ahora un campo $\mathbf{F} = (F_1, F_2, \ldots, F_n) \in \Phi(U)$. Este se puede identificar según lo presentado anteriormente con la $(n-1)$-forma en U

$$\eta = \sum_{i=1}^{n} (-1)^{i+1} F_i dx_1 dx_2 \cdots dx_{i-1} dx_{i+1} \cdots dx_n$$

Siendo η una $(n-1)$-forma, su diferencial exterior será una n-forma en $U \subseteq \mathbb{R}^n$. Obtengámosla

$$\begin{aligned} d\eta &= \sum_{i=1}^{n} (-1)^{i+1} d(F_i) dx_1 dx_2 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \\ &= \sum_{i=1}^{n} (-1)^{i+1} \left(\sum_{j=1}^{n} \frac{\partial F_i}{\partial x_j} dx_j \right) dx_1 dx_2 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \end{aligned}$$

Nótese que de los n^n términos en la expresión anterior, solamente los correspondientes a $\sum_{j=1}^{n} \frac{\partial F_i}{\partial x_j} dx_j$ *cuando* $j = i$, producirán términos no nulos. Se tiene entonces

$$d\eta = \sum_{i=1}^{n} (-1)^{i+1} \frac{\partial F_i}{\partial x_i} dx_i dx_1 dx_2 \cdots dx_{i-1} dx_{i+1} \cdots dx_n$$

Para llevar a dx_i "al lugar que le corresponde", habrá que pasar por las i - 1 expresiones $dx_1, dx_2, \ldots, dx_{i-1}$. Como sabemos, cada uno de estos intercambios produce un signo menos en la expresión anterior, teniendo entonces que

$$\begin{aligned} d\eta &= \sum_{i=1}^{n} (-1)^{i+1} (-1)^{i-1} \frac{\partial F_i}{\partial x_i} dx_1 dx_2 \cdots dx_{i-1} dx_i dx_{i+1} \cdots dx_n \\ &= \left(\sum_{i=1}^{n} \frac{\partial F_i}{\partial x_i} \right) dx_1 dx_2 \cdots dx_n \end{aligned}$$

El lector quedará gratamente sorprendido al identificar en la expresión anterior la divergencia del campo $\mathbf{F}$. Se tiene entonces que $d\eta$ es la n-forma en U de $\mathbb{R}^n$ dada por

$$d\eta = \operatorname{div} \mathbf{F} dx_1 dx_2 \cdots dx_n$$

la cual a su vez queda identificada con la 0-forma div $\mathbf{F}$. Así pues, *la divergencia del campo vectorial* **F** *se identifica con la diferencial exterior de la forma que le corresponde a* **F** *según el isomorfismo* $T_{1\,n-1}$.

Ejemplo 8. Consideremos el campo vectorial en $\mathbb{R}^3$

$$\mathbf{F}(x_1, x_2, x_3) = (3x_1x_2, \operatorname{sen}(x_1x_3), x_2x_3^3)$$

Este se identifica, bajo el isomorfismo $T_{1,2}$, con la 2-forma en $\mathbb{R}^3$

$$\omega = 3x_1x_2dx_2dx_3 - \operatorname{sen}(x_1x_3)dx_1dx_3 + x_2x_3^3dx_1dx_2$$

cuya diferencial exterior es, según lo discutido anteriormente expuesto, la 3-forma en $\mathbb{R}^3$

$$d\omega = \operatorname{div} \mathbf{F}dx_1dx_2dx_3 = (3x_2 + 3x_2x_3^2)dx_1dx_2dx_3$$ ■

Consideremos por último, el campo $\mathbf{F}: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F} = (F_1, F_2, F_3)$. A éste le corresponde la 1-forma en U, $\omega = F_1dx_1 + F_2dx_2 + F_3dx_3$. La diferencial exterior de esta forma es

$$\begin{aligned}
d\omega &= d(F_1)dx_1 + d(F_2)dx_2 + d(F_3)dx_3 \\
&= \left(\frac{\partial F_1}{\partial x_1}dx_1 + \frac{\partial F_1}{\partial x_2}dx_2 + \frac{\partial F_1}{\partial x_3}dx_3\right)dx_1 + \left(\frac{\partial F_2}{\partial x_1}dx_1 + \frac{\partial F_2}{\partial x_2}dx_2 + \frac{\partial F_2}{\partial x_3}dx_3\right)dx_2 \\
&\quad + \left(\frac{\partial F_3}{\partial x_1}dx_1 + \frac{\partial F_3}{\partial x_2}dx_2 + \frac{\partial F_3}{\partial x_3}dx_3\right)dx_3 \\
&= \frac{\partial F_1}{\partial x_2}dx_2dx_1 + \frac{\partial F_1}{\partial x_3}dx_3dx_1 + \frac{\partial F_2}{\partial x_1}dx_1dx_2 + \frac{\partial F_2}{\partial x_3}dx_3dx_2 \\
&\quad + \frac{\partial F_3}{\partial x_1}dx_1dx_3 + \frac{\partial F_3}{\partial x_2}dx_2dx_3 \\
&= -\frac{\partial F_1}{\partial x_2}dx_1dx_2 - \frac{\partial F_1}{\partial x_3}dx_1dx_3 + \frac{\partial F_2}{\partial x_1}dx_1dx_2 - \frac{\partial F_2}{\partial x_3}dx_2dx_3 \\
&\quad + \frac{\partial F_3}{\partial x_1}dx_1dx_3 + \frac{\partial F_3}{\partial x_2}dx_2dx_3 \\
&= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}\right)dx_2dx_3 - \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}\right)dx_1dx_3 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right)dx_1dx_2
\end{aligned}$$

Nótese que ésta es precisamente la 2-forma que le corresponde, bajo el isomorfismo $T_{1,2}$ al campo

$$\operatorname{rot} \mathbf{F} = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right)$$

Así pues, *el rotacional de un campo* **F** *se identifica con la diferencial exterior de la* 1*-forma que le corresponde a* **F**.

Ejemplo 9. Consideremos el campo vectorial en $\mathbb{R}^3$

$$\mathbf{F}(x_1, x_2, x_3) = (5x_1 + 2x_2x_3, x_2x_3, x_1 + x_2 - 3x_3)$$

La 1-forma que le corresponde a **F** es

$$\omega = (5x_1 + 2x_2x_3)dx_1 + x_2x_3dx_2 + (x_1 + x_2 - 3x_3)dx_3$$

cuya diferencial exterior es, según lo previamente expuesto, la 2-forma en $\mathbb{R}^3$ que es identificada con el campo rot **F**, según el isomorfismo $T_{1,2}$. Puesto que

$$\text{rot}\,\mathbf{F} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ 5x_1 + 2x_2x_3 & x_2x_3 & x_1 + x_2 - 3x_3 \end{bmatrix} = (1 - x_2, 2x_2 - 1, -2x_3)$$

la diferencial exterior de la 1-forma ω, debe ser la 2-forma

$$d\omega = (1 - x_2)dx_2dx_3 - (2x_2 - 1)dx_1dx_3 - 2x_3dx_1dx_2$$ ■

Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una 0-forma en U. La diferencial exterior de ella es la 1-forma

$$\omega = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n$$

la cual está identificada con el campo grad f. Por otra parte, el rotacional del campo grad f se identifica con la diferencial exterior de ω. Así, el rotacional del campo grad f es la diferencial exterior de ω, la cual a su vez es la diferencial exterior de la 0-forma f. En términos de diferenciales exteriores, el rotacional del campo grad f queda escrito como $d(df)$. Pero, como sabemos, $d(df) = 0$ (teorema 10.2.3). Es decir, rot(grad f) $= 0$, que es la fórmula 8 de la sección 6 del capítulo 9, ya verificada en forma directa.

Por otra parte, sea $\mathbf{F}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ un campo vectorial en U. El rotacional de **F** es un campo vectorial que se identifica con una 2-forma ω, según el isomorfismo $T_{1,2}$, la cual resulta ser la diferencial exterior de la 1-forma que le corresponde a **F**. A su vez, la diferencial exterior de ω se identifica con la divergencia del campo rot F. En resumen, la divergencia del rotacional de **F** se identifica con la diferencial exterior de la diferencial exterior de la 1-forma correspondiente a **F**. De nuevo, por el teorema 10.2.3, ésta es cero. Es decir, div rot **F** $= 0$, que es la fórmula 9 de la sección 6 del capítulo 9.

Ejercicios (Capítulo 10, Sección 2)

En los ejercicios 1–10, obtenga la diferencial exterior de la forma.

1. $\omega = dx_1 + dx_2$

2. $\omega = x_1x_2x_3dx_1 + x_2dx_2 + dx_3$

3. $\omega = x_4dx_1 + x_3dx_2 + x_2dx_3 + x_1dx_4$

4. $\omega = x_1^2x_2dx_1dx_4 + \operatorname{sen} x_3dx_2dx_3$

5. $\omega = x_1 \operatorname{sen} x_2dx_2dx_4 + (x_2x_5 + x_3)dx_3dx_5 + (x_1x_3 + x_2x_4x_5)dx_4dx_5$

6. $\omega = x_1^2x_2^2x_3^2dx_1dx_2dx_3 + x_1x_4^3dx_2dx_3dx_4$

7. $\omega = (x_1x_2x_3 + x_1x_2x_3 \operatorname{sen}(x_1^2x_2^2x_3^2) + x_1x_2x_3 \cos(x_1^3x_2^3x_3^3))dx_1dx_2dx_3$

8. $\omega = x_2 \operatorname{sen} x_3 \cos x_4dx_2dx_3dx_5 + x_3^3dx_1dx_3dx_5$

9. $\omega = x_1 \ln x_2 dx_1 dx_2 dx_5 dx_6 + x_2 x_3 x_6^3 dx_3 dx_4 dx_5 dx_6$

10. $\omega = x_2 x_3 x_4 \cos^2(x_1 x_2) dx_1 dx_5 dx_7 + x_1 x_3 x_6 \operatorname{sen}^2(x_1 + x_3 x_6) dx_3 dx_6 dx_7$

11. Verifique la fórmula del teorema 10.2.2 con las formas $\omega = x_1 x_2 dx_1 + x_2 x_3 dx_2 + x_3 dx_3$, $\eta = x_3 dx_1 + x_2 dx_2 + x_1 dx_3$.

12. Sean $\omega = x_1 x_2 x_3 dx_1 dx_2 + x_4 dx_2 dx_4$, $\eta = x_1 dx_1 + x_3 dx_3$, $\rho = x_1 x_2 x_3 dx_1$. Verifique la generalización de la fórmula del teorema 10.2.2 con estas formas. Es decir, compruebe que

$$d(\omega\eta\rho) = (d\omega)\eta\rho + (-1)^p \omega d\eta\rho + (-1)^{p+q} \omega\eta d\rho$$

(en el que $p = 2$, $q = 1$).

13. Considere el campo $\mathbf{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{F}(x, y, z) = (x^2 y, xz, yz^2)$.
    a. Determine la 1-forma ω en $\mathbb{R}^3$ con que se identifica $\mathbf{F}$.
    b. Determine la diferencial exterior de ω.
    c. Halle el rotacional de $\mathbf{F}$.
    d. Determine la 2-forma en $\mathbb{R}^3$ con que se identifica la diferencial exterior de ω obtenida en el inciso **b**.

14. Considere el campo $\mathbf{F}: \mathbb{R}^4 \to \mathbb{R}^4$, $\mathbf{F}(x_1, x_2, x_3, x_4) = (x_1 x_4, x_2 x_3, x_4, x_3)$. Este se identifica con una 1-forma ω en $\mathbb{R}^4$ y con una 3-forma η en $\mathbb{R}^4$. Determine estas formas. Determine la divergencia de $\mathbf{F}$ y la diferencial exterior de η.

15. El campo $\mathbf{F}: \mathbb{R}^5 \to \mathbb{R}^5$, $\mathbf{F}(x_1, x_2, x_3, x_4, x_5) = (x_5, x_4, x_3, x_2, x_1)$ se identifica con una 1-forma ω en $\mathbb{R}^5$ y con una 4-forma η en $\mathbb{R}^5$. Determine estas formas. Determine la divergencia de $\mathbf{F}$ y la diferencial exterior de η.

10.3 Cambio de variables en formas

El problema que motiva el estudio de la presente sección es el siguiente: sea ω una p-forma definida en el conjunto abierto U de $\mathbb{R}^n$; se trata de asociar a la forma ω una nueva p-forma $\tilde{\omega}$ definida en *otro conjunto V del espacio* $\mathbb{R}^m$ (con m no necesariamente igual a n). Si llamamos $(x_1, x_2, \ldots, x_n)$ a los vectores de U, y $(y_1, y_2, \ldots, y_m)$ a los vectores de V, podemos ver el problema anterior como un problema de "cambio de variables" en la forma ω: las variables de ésta son $x_1, x_2, \ldots, x_n$, y queremos cambiar estas variables por $y_1, y_2, \ldots, y_m$, obteniendo así la nueva forma $\tilde{\omega}$. Esta nueva operación en las formas es de fundamental importancia para la teoría de integración de formas que comenzaremos a desarrollar en la sección siguiente.

Entremos en los detalles.

Sea entonces U un conjunto abierto del espacio $\mathbb{R}^n$ y V un conjunto abierto del espacio $\mathbb{R}^m$. Denotaremos con x's a las coordenadas de $\mathbb{R}^n$, y con y's a las coordenadas de $\mathbb{R}^m$. Sea $\varphi: V \to U$ una función de clase $\mathscr{C}^2$, $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$ (donde, entonces $\varphi_i: V \to \mathbb{R}$ es una función de clase $\mathscr{C}^2$, $i = 1, 2, \ldots, n$). Esta función φ tendrá el papel de "función de cambio de variables", según lo comentado en el párrafo anterior.

Sea ω una p-forma en U. Asociaremos a ésta una nueva p-forma en V que se denotará $\varphi^*\omega$ (leemos "fi estrella omega"), y que llamaremos *forma inducida (por φ)*. A esta operación en las formas se le suele llamar "operación estrella", o bien, "pull-back" en inglés.

Comencemos por considerar el caso $p = 0$. Si $\omega = f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una 0-forma en U, definimos la 0-forma $\varphi^*\omega$ en V como

$$\varphi^*\omega = f \circ \varphi$$

Así la operación de llevar 0-formas en U a 0-formas en V por medio de la función φ, es una simple composición de la 0-forma f en U con φ.

Para la 1-forma $\omega = dx_i$ en U (i un índice fijo entre 1 y n), definimos la 1-forma $\varphi^*\omega$ en V como

$$\varphi^*(dx_i) = \frac{\partial \varphi_i}{\partial y_1} dy_1 + \frac{\partial \varphi_i}{\partial y_2} dy_2 + \cdots + \frac{\partial \varphi_i}{\partial y_m} dy_m$$

Para una 1-forma ω en U, digamos

$$\omega = g_1 dx_1 + g_2 dx_2 + \cdots + g_n dx_n$$

definimos la 1-forma $\varphi^*\omega$ en V como

$$\varphi^*\omega = \varphi^*(g_1)\varphi^*(dx_1) + \varphi^*(g_2)\varphi^*(dx_2) + \cdots + \varphi^*(g_n)\varphi^*(dx_n)$$

o bien, de modo más explícito

$$\begin{aligned} \varphi^*\omega &= (g_1 \circ \varphi) \sum_{j=1}^{m} \frac{\partial \varphi_1}{\partial y_j} dy_j + (g_2 \circ \varphi) \sum_{j=1}^{m} \frac{\partial \varphi_2}{\partial y_j} dy_j + \cdots + (g_n \circ \varphi) \sum_{j=1}^{m} \frac{\partial \varphi_n}{\partial y_j} dy_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} (g_i \circ \varphi) \frac{\partial \varphi_i}{\partial y_j} dy_j \end{aligned}$$

En general, si ω es una p-forma en U, digamos

$$\omega = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

La p-forma $\varphi^*\omega$ en V se define como

$$\varphi^*\omega = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} \varphi^*(g_{i_1 i_2 \ldots i_p}) \varphi^*(dx_{i_1}) \varphi^*(dx_{i_2}) \cdots \varphi^*(dx_{i_p})$$

Deseamos llamar la atención a cómo se efectúa el cambio de las variables x's de la forma ω por las nuevas variables y's de la forma $\varphi^*\omega$. Tomemos la p-forma $\omega = g dx_1 dx_2 \cdots dx_p$ definida en el abierto $U \subseteq \mathbb{R}^n$, y sea $\varphi: V \subseteq \mathbb{R}^m \to U \subseteq \mathbb{R}^n$ la función de clase $\mathscr{C}^2$ que se encargará de cambiar las variables en ω. Si $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$, $(y_1, y_2, \ldots, y_m) \in V$, $(x_1, x_2, \ldots, x_n) \in U$, podemos escribir entonces

$$x_i = \varphi_i(y_1, y_2, \ldots, y_m)$$

$i = 1, 2, \ldots, n$. El proceso de obtención de la forma inducida $\varphi^*\omega$, se puede ver como un verdadero "cambio de variables": cada x_i se sustituye por la función $\varphi_i(y_1, y_2, \ldots, y_m)$ (lo cual se puede ver globalmente como *componer* la función $g(x_1, x_2, \ldots, x_n)$ con la función $\varphi(y_1, y_2, \ldots, y_m)$), y cada

dx_i lo sustituimos por *la diferencial* de la función real de m variables $x_i = \varphi_i(y_1, y_2, \ldots, y_m)$, es decir por

$$dx_i = \frac{\partial \varphi_i}{\partial y_1} dy_1 + \frac{\partial \varphi_i}{\partial y_2} dy_2 + \cdots + \frac{\partial \varphi_i}{\partial y_m} dy_m$$

(obsérvese que esta fórmula es con la que se definió a $\varphi^*(dx_i)$).

Veamos algunos ejemplos.

Ejemplo 1. Sea ω la 1-forma en $\mathbb{R}$, $\omega = 2x_1^2 dx_1$, y sea $\varphi: \mathbb{R}^2 \to \mathbb{R}$, la función $\varphi(y_1, y_2) = 3y_1 - 4y_2$. La forma $\varphi^*\omega$ será entonces una 1-forma en $\mathbb{R}^2$, a saber

$$\begin{aligned}\varphi^*\omega &= \varphi^*(2x_1^2)\varphi^*(dx_1) = 2(3y_1 - 4y_2)^2\left(\frac{\partial \varphi}{\partial y_1} dy_1 + \frac{\partial \varphi}{\partial y_2} dy_2\right) \\ &= 2(3y_1 - 4y_2)^2(3dy_1 - 4dy_2) = 6(3y_1 - 4y_2)^2 dy_1 - 8(3y_1 - 4y_2)^2 dy_2\end{aligned}$$

■

Ejemplo 2. Sea ω la 1-forma en $\mathbb{R}^2$, $\omega = 2x_1x_2dx_1 + x_1dx_2$, y sea $\varphi: \mathbb{R}^3 \to \mathbb{R}^2$, la función dada por $\varphi(y_1, y_2, y_3) = (2y_2y_3, 3y_1y_2) = (x_1, x_2)$. La forma $\varphi^*\omega$ será entonces una 1-forma en $\mathbb{R}^3$, que se determina a continuación

$$\begin{aligned}\varphi^*\omega &= \varphi^*(2x_1x_2)\varphi^*(dx_1) + \varphi^*(x_1)\varphi^*(dx_2) \\ &= 2(2y_2y_3)(3y_1y_2)\left(\frac{\partial \varphi_1}{\partial y_1} dy_1 + \frac{\partial \varphi_1}{\partial y_2} dy_2 + \frac{\partial \varphi_1}{\partial y_3} dy_3\right) \\ &+ 2y_2y_3\left(\frac{\partial \varphi_2}{\partial y_1} dy_1 + \frac{\partial \varphi_2}{\partial y_2} dy_2 + \frac{\partial \varphi_2}{\partial y_3} dy_3\right) \\ &= 12y_1y_2^2y_3(2y_3dy_2 + 2y_2dy_3) + 2y_2y_3(3y_2dy_1 + 3y_1dy_2) \\ &= 6y_2^2y_3dy_1 + (24y_1y_2^2y_3^2 + 6y_1y_2y_3)dy_2 + 24y_1y_2^3y_3dy_3\end{aligned}$$

■

Ejemplo 3. Sea ω la 2-forma en $\mathbb{R}^3$, $\omega = 5x_1x_2x_3dx_2dx_3$, y sea $\varphi: \mathbb{R}^2 \to \mathbb{R}^3$, la función dada por $\varphi(y_1, y_2) = (3y_1y_2, 5y_1^2, 4y_2^3) = (x_1, x_2, x_3)$. Se tiene

$$\begin{aligned}\varphi^*\omega &= \varphi^*(5x_1x_2x_3)\varphi^*(dx_2)\varphi^*(dx_3) \\ &= 5(3y_1y_2)(5y_1^2)(4y_2^3)\left(\frac{\partial \varphi_2}{\partial y_1} dy_1\right)\left(\frac{\partial \varphi_3}{\partial y_1} dy_1\right) \\ &= 300y_1^3y_2^4(10y_1dy_1)(12y_2^2dy_2) = 36000y_1^4y_2^6dy_1dy_2\end{aligned}$$

la cual es una 2-forma en $\mathbb{R}^2$. ■

Ejemplo 4. Consideremos las 1-formas ω en $\mathbb{R}^2$ y $\varphi^*\omega$ en $\mathbb{R}^3$ del ejemplo 2. Obtengamos sus diferenciales exteriores. La diferencial exterior de ω es

$$\begin{aligned}d\omega &= d(2x_1x_2)dx_1 + d(x_1)dx_2 \\ &= (2x_2dx_1 + 2x_1dx_2)dx_1 + dx_1dx_2 = (1 - 2x_1)dx_1dx_2\end{aligned}$$

La diferencial exterior de $\varphi^*\omega$ es

$$\begin{aligned}
d(\varphi^*\omega) &= d(6y_2^2y_3)dy_1 + d(24y_1y_2^2y_3^2 + 6y_1y_2y_3)dy_2 + d(24y_1y_2^3y_3)dy_3 \\
&= (12y_2y_3dy_2 + 6y_2^2dy_3)dy_1 \\
&\quad + \big((24y_2^2y_3^2 + 6y_2y_3)dy_1 + (48y_1y_2y_3^2 + 6y_1y_3)dy_2 + (48y_1y_2^2y_3 + 6y_1y_2)dy_3\big)dy_2 \\
&\quad + (24y_2^3y_3dy_1 + 72y_1y_2^2y_3dy_2 + 24y_1y_2^3dy_3)dy_3 \\
&= (24y_2^2y_3^2 - 6y_2y_3)dy_1dy_2 + (24y_2^3y_3 - 6y_2^2)dy_1dy_3 + (24y_1y_2^2y_3 - 6y_1y_2)dy_2dy_3
\end{aligned}$$

Hagamos ahora la operación estrella con la función φ y con la diferencial exterior de ω. Se tiene

$$\begin{aligned}
\varphi^*(d\omega) &= \varphi^*(1 - 2x_1)\varphi^*(dx_1)\varphi^*(dx_2) \\
&= \big(1 - 2(2y_2y_3)\big)\left(\frac{\partial\varphi_1}{\partial y_1}dy_1 + \frac{\partial\varphi_1}{\partial y_2}dy_2 + \frac{\partial\varphi_1}{\partial y_3}dy_3\right)\left(\frac{\partial\varphi_2}{\partial y_1}dy_1 + \frac{\partial\varphi_2}{\partial y_2}dy_2 + \frac{\partial\varphi_2}{\partial y_3}dy_3\right) \\
&= (1 - 4y_2y_3)(-6y_2y_3dy_1dy_2 - 6y_2^2dy_1dy_3 - 6y_1y_2dy_2dy_3) \\
&= (24y_2^2y_3^2 - 6y_2y_3)dy_1dy_2 + (24y_2^3y_3 - 6y_2^2)dy_1dy_3 + (24y_1y_2^2y_3 - 6y_1y_2)dy_2dy_3 \\
&= d(\varphi^*\omega)
\end{aligned}$$

Se ha verificado entonces, en este ejemplo particular, que $\varphi^*(d\omega) = d(\varphi^*\omega)$. Esta es una propiedad general que se estudiará posteriormente. ■

Comencemos ahora el estudio de las propiedades de la operación estrella en formas.

Teorema 10.3.1 Sean ω y ω' dos p-formas en el conjunto abierto U de $\mathbb{R}^n$, η una q-forma en U, y $\varphi\colon V \to U$ una función de clase $\mathscr{C}^2$ en el conjunto abierto V de $\mathbb{R}^m$. Entonces

$$\varphi^*(\omega + \omega') = \varphi^*\omega + \varphi^*\omega'$$
$$\varphi^*(\omega\eta) = (\varphi^*\omega)(\varphi^*\eta)$$

Demostración. Se trata de consecuencias inmediatas de las definiciones correspondientes de suma y producto de formas, y de la operación estrella en éstas. Presentamos las operaciones que validan la segunda de las fórmulas del teorema y dejamos como ejercicio para el lector la primera de ellas. Sean entonces

$$\omega = \sum_{1\le i_1<i_2<\cdots<i_p\le n} g_{i_1i_2\ldots i_p}dx_{i_1}dx_{i_2}\cdots dx_{i_p}$$
$$\eta = \sum_{1\le j_1<j_2<\cdots<j_q\le n} h_{j_1j_2\ldots j_q}dx_{j_1}dx_{j_2}\cdots dx_{j_q}$$

El producto de estas formas es

$$\omega\eta = \sum_{\substack{1\le i_1<i_2<\cdots<i_p\le n\\ 1\le j_1<j_2<\cdots<j_q\le n}} g_{i_1i_2\ldots i_p}h_{j_1j_2\ldots j_q}dx_{i_1}\cdots dx_{i_p}dx_{j_1}\cdots dx_{j_q}$$

de modo que

$$\varphi^*(\omega\eta) = \sum_{\substack{1\le i_1<i_2<\cdots<i_p\le n\\ 1\le j_1<j_2<\cdots<j_q\le n}} \varphi^*(g_{i_1i_2\ldots i_p}h_{j_1j_2\ldots j_q})\varphi^*(dx_{i_1})\cdots\varphi^*(dx_{i_p})\varphi^*(dx_{j_1})\cdots\varphi^*(dx_{j_q})$$

Nótese que

$$\varphi^*(g_{i_1i_2\ldots i_p}h_{j_1j_2\ldots j_q}) = (g_{i_1i_2\ldots i_p}h_{j_1j_2\ldots j_q})\circ\varphi$$
$$= (g_{i_1i_2\ldots i_p}\circ\varphi)(h_{j_1j_2\ldots j_q}\circ\varphi) = \varphi^*(g_{i_1i_2\ldots i_p})\varphi^*(h_{j_1j_2\ldots j_q})$$

Entonces

$$\varphi^*(\omega\eta) = \sum_{\substack{1\le i_1<i_2<\cdots<i_p\le n\\ 1\le j_1<j_2<\cdots<j_q\le n}} \varphi^*(g_{i_1i_2\ldots i_p})\varphi^*(h_{j_1j_2\ldots j_q})\varphi^*(dx_{i_1})\cdots\varphi^*(dx_{i_p})\varphi^*(dx_{j_1})\ldots\varphi^*(dx_{j_q})$$
$$= \left(\sum_{1\le i_1<i_2<\cdots<i_p\le n}\varphi^*(g_{i_1i_2\ldots i_p})\varphi^*(dx_{i_1})\cdots\varphi^*(dx_{i_p})\right)\left(\sum_{1\le j_1<j_2<\cdots<j_q\le n}\varphi^*(h_{j_1j_2\ldots j_p})\varphi^*(dx_{j_1})\cdots\varphi^*(dx_{j_p})\right)$$
$$= (\varphi^*\omega)(\varphi^*\eta)$$

Q.E.D

Se ha visto entonces cómo llevar p-formas ω definidas en el conjunto abierto U de $\mathbb{R}^n$, a p-formas $\varphi^*\omega$ definidas en el conjunto abierto V de $\mathbb{R}^m$, por medio de una función $\varphi\colon V\to U$, problema que, habíamos dicho ya, se puede contemplar como el cambio de variables x's de ω por variables y's en $\varphi^*\omega$. Supongamos ahora que se quiere llevar la p-forma $\varphi^*\omega$ en $V\subseteq\mathbb{R}^m$, a una p-forma $\psi^*(\varphi^*\omega)$ definida en el conjunto abierto W de $\mathbb{R}^k$, por medio de la función $\psi\colon W\to V$. Llamando z's a las coordenadas de los vectores de W, se intenta cambiar ahora las coordenadas y's de $\varphi^*\omega$ por coordenadas z's. En el siguiente teorema se recoge el comportamiento de las operaciones estrella corrrespondientes a φ y ψ en la forma ω.

Teorema 10.3.2 Sea ω una p-forma definida en el conjunto abierto $U\subseteq\mathbb{R}^n$. Sea $\varphi\colon V\to U$ una función de clase $\mathscr{C}^2$ definida en el conjunto abierto $V\subseteq\mathbb{R}^m$, y $\psi\colon W\to V$ una función de clase $\mathscr{C}^2$ definida en el conjunto abierto $W\subseteq\mathbb{R}^k$. Entonces

$$(\varphi\circ\psi)^*\omega = \psi^*(\varphi^*\omega)$$

o bien

$$(\varphi\circ\psi)^* = \psi^*\circ\varphi^*$$

Demostración. Llamaremos entonces x's a las coordenadas de los vectores de U, y's a las coordenadas de los vectores de V, y z's a las coordenadas de los vectores de W. Escribiremos también $\varphi = (\varphi_1,\varphi_2,\ldots,\varphi_n)$ y $\psi = (\psi_1,\psi_2,\ldots,\psi_m)$. Obsérvese que $\varphi\circ\psi = (\varphi_1\circ\psi,\varphi_2\circ\psi,\ldots,\varphi_n\circ\psi)$. Veamos primeramente la validez de la fórmula del teorema, en caso de que ω sea una 0-forma en $f\colon U\subseteq\mathbb{R}^n\to\mathbb{R}$. Se tiene

$$(\varphi\circ\psi)^*f = f\circ(\varphi\circ\psi) = (f\circ\varphi)\circ\psi = \psi^*(f\circ\varphi) = \psi^*(\varphi^*f)$$

Consideremos ahora el caso en que ω en la 1-forma dx_i (i un índice fijo entre 1 y n). Se tiene

$$(\varphi\circ\psi)^*(dx_i) = \frac{\partial}{\partial z_1}(\varphi_i\circ\psi)dz_1 + \frac{\partial}{\partial z_2}(\varphi_i\circ\psi)dz_2 + \cdots + \frac{\partial}{\partial z_k}(\varphi_i\circ\psi)dz_k$$
$$= \sum_{j=1}^{k}\frac{\partial}{\partial z_j}(\varphi_i\circ\psi)dz_j$$

Por la regla de la cadena se tiene, escribiendo $z = (z_1, z_2, \ldots, z_k) \in W$

$$\begin{aligned}\frac{\partial}{\partial z_j}(\varphi_i \circ \psi)(z) &= \frac{\partial \varphi_i}{\partial y_1}(\psi(z))\frac{\partial \psi_1}{\partial z_j}(z) + \cdots + \frac{\partial \varphi_i}{\partial y_m}(\psi(z))\frac{\partial \psi_m}{\partial z_j}(z) \\ &= \sum_{t=1}^{m} \frac{\partial \varphi_i}{\partial y_t}(\psi(z))\frac{\partial \psi_t}{\partial z_j}(z)\end{aligned}$$

que podemos escribir como

$$\frac{\partial}{\partial z_j}(\varphi_i \circ \psi) = \sum_{t=1}^{m} \frac{\partial \varphi_i}{\partial y_t} \circ \psi \frac{\partial \psi_t}{\partial z_j}$$

Entonces

$$\begin{aligned}(\varphi \circ \psi)^*(dx_i) &= \sum_{j=1}^{k} \frac{\partial}{\partial z_j}(\varphi_i \circ \psi)dz_j = \sum_{j=1}^{k}\sum_{t=1}^{m} \frac{\partial \varphi_i}{\partial y_t} \circ \psi \frac{\partial \psi_t}{\partial z_j}dz_j \\ &= \sum_{t=1}^{m} \frac{\partial \varphi_i}{\partial y_t} \circ \psi \sum_{j=1}^{k} \frac{\partial \psi_t}{\partial z_j}dz_j = \sum_{t=1}^{m} \psi^*\left(\frac{\partial \varphi_i}{\partial y_t}\right)\psi^*(dy_t) = \psi^* \sum_{t=1}^{m} \frac{\partial \varphi_i}{\partial y_t}dy_t \\ &= \psi^*(\varphi^*(dx_i))\end{aligned}$$

Sea ahora ω una p-forma cualquiera en U, digamos

$$\omega = \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}$$

Usando los resultados previos, se tiene

$$\begin{aligned}(\varphi \circ \psi)^*\omega &= \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} (\varphi \circ \psi)^*(g_{i_1 i_2 \ldots i_p})(\varphi \circ \psi)^*(dx_{i_1})(\varphi \circ \psi)^*(dx_{i_2}) \cdots (\varphi \circ \psi)^*(dx_{i_p}) \\ &= \sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} \psi^*(\varphi^*(g_{i_1 i_2 \ldots i_p}))\psi^*(\varphi^*(dx_{i_1}))\psi^*(\varphi^*(dx_{i_2})) \cdots \psi^*(\varphi^*(dx_{i_p})) \\ &= \psi^*\left(\sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} \varphi^*(g_{i_1 i_2 \ldots i_p})\varphi^*(dx_{i_1})\varphi^*(dx_{i_2}) \cdots \varphi^*(dx_{i_p})\right) \\ &= \psi^*\left(\varphi^*\left(\sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} g_{i_1 i_2 \ldots i_p} dx_{i_1} dx_{i_2} \cdots dx_{i_p}\right)\right) \\ &= \psi^*(\varphi^*\omega)\end{aligned}$$

Q.E.D.

Ejemplo 5. Sea ω la 1-forma en $\mathbb{R}^2$, $\omega = 2x_1x_2dx_1 + x_1dx_2$, y sea $\varphi: \mathbb{R}^3 \to \mathbb{R}^2$, la función dada por $\varphi(y_1, y_2, y_3) = (2y_2y_3, 3y_1y_2) = (x_1, x_2)$. En el ejemplo 2 se obtuvo que

$$\varphi^*\omega = 6y_2^2y_3dy_1 + (24y_1y_2^2y_3^2 + 6y_1y_2y_3)dy_2 + 24y_1y_2^3y_3dy_3$$

Consideremos la función $\psi: \mathbb{R}^4 \to \mathbb{R}^3$, $\psi = (\psi_1, \psi_2, \psi_3)$, dada por

$$\psi(z_1, z_2, z_3, z_4) = (3z_2, z_1z_4, z_2z_3) = (y_1, y_2, y_3)$$

Obtengamos $\psi^*(\varphi^*\omega)$. Se tiene

$$\begin{aligned}
\psi^*(\varphi^*\omega) &= \psi^*(6y_2^2y_3dy_1 + (24y_1y_2^2y_3^2 + 6y_1y_2y_3)dy_2 + 24y_1y_2^3y_3dy_3) \\
&= \psi^*(6y_2^2y_3)\psi^*(dy_1) + \psi^*(24y_1y_2^2y_3^2 + 6y_1y_2y_3)\psi^*(dy_2) + \psi^*(24y_1y_2^3y_3)\psi^*(dy_3) \\
&= 6(z_1z_4)^2z_2z_3\left(\frac{\partial\psi_1}{\partial z_1}dz_1 + \frac{\partial\psi_1}{\partial z_2}dz_2 + \frac{\partial\psi_1}{\partial z_3}dz_3 + \frac{\partial\psi_1}{\partial z_4}dz_4\right) \\
&\quad + \left(24(3z_2)(z_1z_4)^2(z_2z_3)^2 + 6(3z_2)(z_1z_4)(z_2z_3)\right)\left(\frac{\partial\psi_2}{\partial z_1}dz_1 + \frac{\partial\psi_2}{\partial z_2}dz_2 + \frac{\partial\psi_2}{\partial z_3}dz_3 + \frac{\partial\psi_2}{\partial z_4}dz_4\right) \\
&\quad + 24(3z_2)(z_1z_4)^3(z_2z_3)\left(\frac{\partial\psi_3}{\partial z_1}dz_1 + \frac{\partial\psi_3}{\partial z_2}dz_2 + \frac{\partial\psi_3}{\partial z_3}dz_3 + \frac{\partial\psi_3}{\partial z_4}dz_4\right) \\
&= 6z_1^2z_2z_3z_4^2(3dz_2) + (72z_1^2z_2^3z_3^2z_4^2 + 18z_1z_2^2z_3z_4)(z_4dz_1 + z_1dz_4) + 72z_1^3z_2^2z_3z_4^3(z_3dz_2 + z_2dz_3) \\
&= (72z_1^2z_2^3z_3^2z_4^3 + 18z_1z_2^2z_3z_4^2)dz_1 + (18z_1^2z_2z_3z_4^2 + 72z_1^3z_2^2z_3^2z_4^3)dz_2 \\
&\quad + 72z_1^3z_2^3z_3z_4^3dz_3 + (72z_1^3z_2^3z_3^2z_4^2 + 18z_1^2z_2^2z_3z_4)dz_4
\end{aligned}$$

Por otra parte, se tiene $\varphi \circ \psi: \mathbb{R}^4 \to \mathbb{R}^2$,

$$\begin{aligned}
(\varphi \circ \psi)(z_1, z_2, z_3, z_4) &= \varphi(\psi(z_1, z_2, z_3, z_4)) = \varphi(3z_2, z_1z_4, z_2z_3) \\
&= (2(z_1z_4)(z_2z_3), 3(3z_2)(z_1z_4)) = (2z_1z_2z_3z_4, 9z_1z_2z_4)
\end{aligned}$$

Obtengamos $(\varphi \circ \psi)^*\omega$. Tenemos

$$\begin{aligned}
(\varphi \circ \psi)^*\omega &= (\varphi \circ \psi)^*(2x_1x_2dx_1 + x_1dx_2) \\
&= (\varphi \circ \psi)^*(2x_1x_2)(\varphi \circ \psi)^*(dx_1) + (\varphi \circ \psi)^*(x_1)(\varphi \circ \psi)^*(dx_2) \\
&= 2(2z_1z_2z_3z_4)(9z_1z_2z_4)\left(2z_2z_3z_4dz_1 + 2z_1z_3z_4dz_2 + 2z_1z_2z_4dz_3 + 2z_1z_2z_3dz_4\right. \\
&\quad \left. + (2z_1z_2z_3z_4)(9z_2z_4dz_1 + 9z_1z_4dz_2 + 9z_1z_2dz_4\right) \\
&= (72z_1^2z_2^3z_3^2z_4^3 + 18z_1z_2^2z_3z_4^2)dz_1 + (18z_1^2z_2z_3z_4^2 + 72z_1^3z_2^2z_3^2z_4^3)dz_2 \\
&\quad + 72z_1^3z_2^3z_3z_4^3dz_3 + (72z_1^3z_2^3z_3^2z_4^2 + 18z_1^2z_2^2z_3z_4)dz_4 = \psi^*(\varphi^*\omega)
\end{aligned}$$

lo que comprueba la fórmula del teorema 10.3.2. ■

Veamos ahora el teorema que establece la propiedad que ya habíamos comentado anteriormente (ver ejemplo 4) que nos dice que los símbolos φ^* y d conmutan. De modo más preciso se tiene el

Teorema 10.3.3 Sea ω una p-forma de clase $\mathscr{C}^2$ definida en el conjunto abierto $U \subseteq \mathbb{R}^n$. Sea $\varphi: V \to U$ una función de clase $\mathscr{C}^2$ definida en el conjunto abierto $V \subseteq \mathbb{R}^m$. Entonces

$$d(\varphi^*\omega) = \varphi^*(d\omega)$$

Demostración. Veamos en primera instancia que la fórmula del teorema se verifique en el caso de una 0-forma, digamos $\omega = f: U \subseteq \mathbb{R}^n \to \mathbb{R}$. Se tiene

$$\begin{aligned}
d(\varphi^*\omega) = d(\varphi^*f) = d(f \circ \varphi) &= \sum_{j=1}^m \frac{\partial}{\partial y_j}(f \circ \varphi)dy_j = \sum_{j=1}^m \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i} \circ \varphi \frac{\partial \varphi_i}{\partial y_j}\right)dy_j \\
&= \sum_{i=1}^n \frac{\partial f}{\partial x_i} \circ \varphi \sum_{j=1}^m \frac{\partial \varphi_i}{\partial y_j}dy_j = \sum_{i=1}^n \varphi^*\frac{\partial f}{\partial x_i}\varphi^*(dx_i) = \varphi^*\sum_{i=1}^n \frac{\partial f}{\partial x_i}dx_i = \varphi^*(df)
\end{aligned}$$

Consideremos ahora la 1-forma $\omega = dx_k$ (k un índice fijo entre 1 y n). Se tiene, por una parte, $\varphi^*(d\omega) = \varphi^*(d(dx_k)) = \varphi^*(0) = 0$, pues $d^2 = 0$. Por otra parte

$$\begin{aligned}
d(\varphi^*\omega) = d(\varphi^*(dx_k)) &= d\left(\sum_{j=1}^{m} \frac{\partial \varphi_k}{\partial y_j} dy_j\right) = \sum_{j=1}^{m} d\left(\frac{\partial \varphi_k}{\partial y_j}\right) dy_j \\
&= \sum_{j=1}^{m}\left(\sum_{i=1}^{m} \frac{\partial}{\partial y_i}\left(\frac{\partial \varphi_k}{\partial y_j}\right) dy_i\right) dy_j = \sum_{i,j=1}^{m} \frac{\partial^2 \varphi_k}{\partial y_i \partial y_j} dy_i dy_j \\
&= \sum_{i<j} \frac{\partial^2 \varphi_k}{\partial y_i \partial y_j} dy_i dy_j + \sum_{i>j} \frac{\partial^2 \varphi_k}{\partial y_i \partial y_j} dy_i dy_j \\
&= \sum_{i<j} \frac{\partial^2 \varphi_k}{\partial y_i \partial y_j} dy_i dy_j + \sum_{j>i} \frac{\partial^2 \varphi_k}{\partial y_j \partial y_i} dy_j dy_i \\
&= \sum_{i<j}\left(\frac{\partial^2 \varphi_k}{\partial y_i \partial y_j} dy_i dy_j + \frac{\partial^2 \varphi_k}{\partial y_j \partial y_i} dy_j dy_i\right) \\
&= \sum_{i<j}\left(\frac{\partial^2 \varphi_k}{\partial y_i \partial y_j} - \frac{\partial^2 \varphi_k}{\partial y_j \partial y_i}\right) dy_i dy_j = 0
\end{aligned}$$

pues la función φ se supone de clase $\mathscr{C}^2$ y entonces sus derivadas parciales de segundo orden cruzadas son iguales (se trata de un argumento similar al presentado en el teorema 10.2.3, en el que se demostró que $d^2 = 0$).

Consideremos ahora el caso de una p-forma del tipo $\omega = f dx_1 dx_2 \cdots dx_p$. Por una parte se tiene que

$$\begin{aligned}
\varphi^*(d\omega) = \varphi^*(df dx_1 dx_2 \cdots dx_p) &= \varphi^*(df)\varphi^*(dx_1)\varphi^*(dx_2)\cdots\varphi^*(dx_p) \\
&= d(\varphi^* f)\varphi^*(dx_1)\varphi^*(dx_2)\cdots\varphi^*(dx_p)
\end{aligned}$$

pues ya probamos que (para la 0-forma f) $\varphi^*(df) = d(\varphi^* f)$. Por otra parte

$$d(\varphi^*\omega) = d\big(\varphi^* f \varphi^*(dx_1)\varphi^*(dx_2)\cdots\varphi^*(dx_p)\big)$$

Esta expresión es la de la diferncial exterior de un producto de $(p+1)$ formas, a saber, la 0-forma $\varphi^* f$ y las 1-formas (p en total) $\varphi^*(dx_1), \varphi^*(dx_2), \ldots, \varphi^*(dx_p)$. Sabemos (de la generalización del teorema 10.2.2) que ésta es igual al producto de la diferencial exterior de la primera forma (en nuestro caso la 0-forma $\varphi^* f$) por el resto de las formas en el producto dado, más una suma de p términos, en cada uno de los cuales aparecerá la diferencial exterior de cada una de las 1-formas $\varphi^*(dx_i)$. Por el caso previamente demostrado, esta diferencial es igual a 0. Así, de los $(p+1)$ sumandos en que se expande la diferencial exterior del producto $\varphi^* f \varphi^*(dx_1)\varphi^*(dx_2)\ldots\varphi^*(dx_p)$ solamente sobrevive el sumando $d(\varphi^* f)\varphi^*(dx_1)\varphi^*(dx_2)\ldots\varphi^*(dx_p)$. Este es justamente el que necesitamos para demostrar que $d(\varphi^*\omega)$ es igual a $\varphi^*(d\omega)$.

Por último, si ω es una p-forma arbitraria, ésta es igual a la suma de términos del tipo $f\,dx_1 dx_2 \cdots dx_p$ (como las consideredas en el caso anterior). Puesto que los símbolos d y φ^* "pasan por las sumas", este caso queda ya demostrado con los casos previos. Q.E.D.

Veremos por último un resultado técnico que usaremos en la sección 6. Sea ω una 2-forma definida en el conjunto abierto $U \subseteq \mathbb{R}^2$, digamos $\omega = f dx_1 dx_2$. Supongamos que $\varphi: V \to U$ una

función de clase $\mathscr{C}^2$, $\varphi = (\varphi_1, \varphi_2)$, definida en el conjunto abierto $V \subseteq \mathbb{R}^2$. Entonces

$$\begin{aligned}\varphi^*\omega &= \varphi^*(fdx_1dx_2) = (f\circ\varphi)\varphi^*(dx_1)\varphi^*(dx_2)\\ &= (f\circ\varphi)\left(\frac{\partial\varphi_1}{\partial y_1}dy_1 + \frac{\partial\varphi_1}{\partial y_2}dy_2\right)\left(\frac{\partial\varphi_2}{\partial y_1}dy_1 + \frac{\partial\varphi_2}{\partial y_2}dy_2\right)\\ &= (f\circ\varphi)\left(\frac{\partial\varphi_1}{\partial y_1}\frac{\partial\varphi_2}{\partial y_2} - \frac{\partial\varphi_2}{\partial y_2}\frac{\partial\varphi_1}{\partial y_2}\right)dy_1dy_2 = (f\circ\varphi)\det\varphi'(y_1, y_2)\end{aligned}$$

Este es un hecho general que se enuncia en el teorema siguiente, cuya demostración se deja como ejercicio para el lector.

Teorema 10.3.4 Sea ω una n-forma definida en el conjunto abierto $U \subseteq \mathbb{R}^n$, digamos $\omega = fdx_1dx_2\cdots dx_n$. Supongamos que $\varphi: V \to U$ una función de clase $\mathscr{C}^2$, definida en el conjunto abierto $V \subseteq \mathbb{R}^n$. Entonces

$$\varphi^*\omega = (f\circ\varphi)\det\varphi'(y)dy_1dy_2\cdots dy_n$$

donde $y = (y_1, y_2, \ldots, y_n) \in V$. ∎

Ejercicios (Capítulo 10, Sección 3)

En los ejercicios 1–10 se dan formas ω definidas en $U \subseteq \mathbb{R}^n$ y funciones $\varphi: V \subseteq \mathbb{R}^m \to U \subseteq \mathbb{R}^n$. Obtenga en cada caso la forma $\varphi^*\omega$.

1. $\omega = x_1dx_1 + x_2dx_2$, $\varphi: \mathbb{R}^2 \to \mathbb{R}^2$, $\varphi(y_1, y_2) = (y_2, y_1)$.

2. $\omega = x_1dx_1 + x_2dx_2$, $\varphi: \mathbb{R}^3 \to \mathbb{R}^2$, $\varphi(y_1, y_2, y_3) = (y_1y_2y_3, y_2y_3)$.

3. $\omega = x_1x_2dx_1dx_2 + x_3dx_2dx_3$, $\varphi: \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(y_1, y_2) = (y_2, y_1, y_1y_2)$.

4. $\omega = x_1x_2x_3dx_1dx_2dx_3$, $\varphi: \mathbb{R}^4 \to \mathbb{R}^3$, $\varphi(y_1, y_2, y_3, y_4) = (y_1^2, y_2y_3, y_3y_4)$.

5. $\omega = x_3 \operatorname{sen} x_1dx_1dx_4 + x_4dx_2dx_4 + \cos x_2dx_3dx_4$, $\varphi: \mathbb{R}^3 \to \mathbb{R}^4$, $\varphi(y_1, y_2, y_3) = (y_1 + y_2 + y_3, y_1 - y_2 - y_3, y_1y_2y_3, y_3)$.

6. $\omega = (x_1 + x_1x_2 + x_1x_2x_3 + x_1x_2x_3x_4)dx_1dx_2dx_4$, $\varphi: \mathbb{R}^2 \to \mathbb{R}^4$, $\varphi(y_1, y_2) = (y_1y_2\cos(y_2), y_1 + y_2\operatorname{sen}(y_1y_2), \cos^2 y_1^2, \operatorname{sen}^2 y_2^2)$. (Sugerencia: piense antes de empezar a hacer operaciones).

7. $\omega = x_1x_2x_3x_4dx_1dx_2dx_3dx_4$, $\varphi: \mathbb{R}^4 \to \mathbb{R}^4$, $\varphi(y_1, y_2, y_3, y_4) = (y_1^2, y_2^2, y_3^2, y_4^2)$

8. $\omega = dx_1dx_2 + dx_2dx_3 - dx_3dx_4$, $\varphi: \mathbb{R}^2 \to \mathbb{R}^4$, $\varphi(y_1, y_2) = (y_1, y_1, y_2, y_2)$.

9. $\omega = x_1x_2\operatorname{sen}(x_3x_4x_5)dx_2dx_3dx_4 + x_2dx_1dx_4dx_5$, $\varphi: \mathbb{R}^4 \to \mathbb{R}^5$, $\varphi(y_1, y_2, y_3, y_4) = (y_1, y_1y_2, y_3, y_4, y_3y_4)$.

10. $\omega = x_1x_3x_5x_7dx_1dx_3dx_5dx_7 + x_2^2x_4^2x_6^2x_8^2dx_2dx_4dx_6dx_8 + x_1^3x_2^3x_7^3x_8^3dx_1dx_2dx_7dx_8$ $\varphi: \mathbb{R}^4 \to \mathbb{R}^8$, $\varphi(y_1, y_2, y_3, y_4) = (1, y_1, 2, y_2, 3, y_3, 4, y_4)$.

11. Considere la 2-forma en $\mathbb{R}^3$, $\omega = x_1x_2dx_1dx_3 + x_3dx_2dx_3$. Sean $\varphi: \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(y_1, y_2) = (y_1+y_2, y_1y_2, y_2)$, $\psi: \mathbb{R}^4 \to \mathbb{R}^2$, $\psi(z_1, z_2, z_3, z_4) = (z_1z_2, z_3z_4)$. Determine: **a.** $\varphi^*\omega$, **b.** $\psi^*(\varphi^*\omega)$, **c.** $(\varphi\circ\psi)^*\omega$. Compruebe que se satisfaga la fórmula del teorema 10.3.2.

12. Sea ω la 3-forma en $\mathbb{R}^5$, $\omega = x_1^2x_2^2x_5^4dx_1dx_4dx_5 + 3x_1^3x_2^4x_4^7dx_2dx_3dx_5$. Sean $\varphi: \mathbb{R}^2 \to \mathbb{R}^5$, $\varphi(y_1, y_2) = (2y_1, y_2^2, y_2^3, y_1^4, y_1^2y_2^6)$, $\psi: \mathbb{R}^2 \to \mathbb{R}^2$, $\psi(z_1, z_2) = (z_1^3 + z_1z_2^4, z_1^6 - 3z_2^4)$. Halle $(\varphi \circ \psi)^*\omega$. (Sugerencia: piense antes de empezar a hacer operaciones).

13. Verifique la fórmula del teorema 10.3.3 con la 2-forma en $\mathbb{R}^4$, $\omega = x_2x_3dx_1dx_4$, y la función $\varphi: \mathbb{R}^3 \to \mathbb{R}^4$, $\varphi(y_1, y_2, y_3) = (y_1y_3, y_2y_4, y_4, y_3)$.

14. Verifique la fórmula del teorema 10.3.4 con la 3-forma en $\mathbb{R}^3$, $\omega = x_1x_2x_3dx_1dx_2dx_3$, y la función $\varphi: \mathbb{R}^3 \to \mathbb{R}^3$, $\varphi(y_1, y_2, y_3) = (y_1 + 2y_2 + 3y_3, 2y_2 + 3y_3, 3y_3)$.

15. Verifique la fórmula del teorema 10.3.4 con la n-forma en $\mathbb{R}^n$,

$$\omega = x_1x_2 \ldots x_ndx_1dx_2 \cdots dx_n$$

y la función $\varphi: \mathbb{R}^n \to \mathbb{R}^n$, $\varphi(y_1, y_2, \ldots, y_n) = (y_1, 2y_2, \ldots, ny_n)$

10.4 Integración de p-formas sobre p-cubos

Sea ω una n-forma definida en el conjunto abierto $U \subseteq \mathbb{R}^n$, digamos $\omega = fdx_1dx_2 \cdots dx_n$, y R una región cerrada y acotada de U, de las consideradas en la sección 9 del capítulo 6 (rectángulos Q en $\mathbb{R}^n$, o regiones limitadas por gráficas de funciones continuas de k variables $k \leq n - 1$). Definimos la integral de ω sobre la región R, denotada por $\int_R \omega$, como la integral n-múltiple de la función $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ sobre R. Es decir

$$\int_R \omega = \int_R f(x_1, x_2, \ldots, x_n)dx_1dx_2 \cdots dx_n$$

Nótese que el lado derecho de esta expresión es una integral ya definida en la sección 9 del capítulo 6.

En el caso general de una p-forma ω definida en un conjunto abierto U de $\mathbb{R}^n$, las regiones en que las integraremos serán unos objetos matemáticos llamados "p-cubos singulares", que presentamos a continuación.

Un p-cubo (unitario) en el espacio $\mathbb{R}^p$ es el conjunto

$$I^p = \{(x_1, x_2, \ldots, x_p)|0 \leq x_i \leq 1, i = 1, 2, \ldots, p\}$$

(ver ejercicio 39 de la sección 3 del capítulo 1). Así, un 1-cubo en $\mathbb{R}$ es el intervalo $[0, 1]$, un 2-cubo en $\mathbb{R}^2$ es el cuadrado con vértices en $(0, 0)$, $(0, 1)$, $(1, 1)$ y $(1, 0)$, un 3-cubo en $\mathbb{R}^3$ es un cubo determinado por el origen de coordenadas y el punto $(1, 1, 1)$ (estos puntos son dos vértices opuestos del cubo), etc. Para el caso $p = 0$, definimos $I^0 = \{0\}$. Un *p-cubo singular* en el conjunto U de $\mathbb{R}^n$ es una función diferenciable $T: I^p \to U \subseteq \mathbb{R}^n$ (o bien, es la restricción al cubo I^p de una función diferenciable $T: V \to U$, definida en el conjunto abierto V de $\mathbb{R}^p$ el cual contiene a I^p). Por ejemplo, un 1-cubo singular en $\mathbb{R}^2$ es una función diferenciable $T: [0, 1] \to \mathbb{R}^2$, la cual, obsérvese, que es una curva en $\mathbb{R}^2$. Un 2-cubo singular en $\mathbb{R}^3$ es una función diferenciable $T: [0, 1] \times [0, 1] \to \mathbb{R}^3$, la cual, obsérvese, puede ser una superficie en $\mathbb{R}^3$.

Sea ω una p-forma definida en el conjunto abierto U de $\mathbb{R}^n$, y $T: I^p \to U$ un p-cubo singular en U. Definimos la *integral de la p-forma ω sobre el p-cubo singular T*, denotada por $\int_T \omega$, como

$$\int_T \omega = \int_{I^p} T^*\omega$$

donde $\int_{I^p} T^*\omega$ es la integral de la p-forma $T^*\omega$ inducida por T, la cual está definida en el conjunto I^p de $\mathbb{R}^p$, sobre el p-cubo I^p, según se definió anteriormente. Si $p = 0$, definimos la integral de la 0-forma $\omega = f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ sobre el 0-cubo singular $T: \{0\} \to U$ como

$$\int_T \omega = \omega(T(0))$$

Veamos, a manera de ejemplos, algunos casos particulares de esta definición.

Ejemplo 1. Consideremos el campo vectorial $F: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $F = (F_1, F_2, \ldots, F_n) \in \Phi(U)$. A este campo asociamos la 1-forma en U

$$\omega = F_1 dx_1 + F_2 dx_2 + \cdots + F_n dx_n$$

Sea $\lambda: [0, 1] \to U$, un 1-cubo singular en U, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Llamemos t a la variable de λ. Se tiene

$$\lambda^*\omega = \lambda^* F_1 \lambda^*(dx_1) + \lambda^* F_2 \lambda^*(dx_2) + \cdots + \lambda^* F_n \lambda^*(dx_n)$$

donde $\lambda^* F_i \doteq F_i \circ \lambda$, y $\lambda^*(dx_i) = \lambda_i' dt$, $i = 1, 2, \ldots, n$. Es decir, la forma $\lambda^*\omega$ (la cual está definida en $[0, 1]$) es

$$\lambda^*\omega(t) = F_1(\lambda(t))\lambda_1'(t)dt + F_2(\lambda(t))\lambda_2'(t)dt + \cdots + F_n(\lambda(t))\lambda_n'(t)dt$$

Entonces, la integral de la 1-forma ω sobre el 1-cubo singular λ es

$$\begin{aligned}\int_\lambda \omega &= \int_{[0,1]} \lambda^*\omega \\ &= \int_0^1 (F_1(\lambda(t))\lambda_1'(t) + F_2(\lambda(t))\lambda_2'(t) + \cdots + F_n(\lambda(t))\lambda_n'(t))dt\end{aligned}$$

la cual no es más que la definición de integral de línea del campo F a lo largo del camino λ. ■

Ejemplo 2. Consideremos el campo vectorial $F: U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, $F = (F_1, F_2, F_3)$. Este campo se puede identificar con la 2-forma ω en U dada por

$$\omega = F_1 dx_2 dx_3 - F_2 dx_1 dx_3 + F_3 dx_1 dx_2$$

(ver parte final de la sección 2). Sea $T: I^2 \to U$ el 2-cubo singular en U, $T = (T_1, T_2, T_3)$, en donde entonces $T_i: [0, 1] \times [0, 1] \to \mathbb{R}$, $i = 1, 2, 3$. Llamemos u y v a las coordenadas de $\mathbb{R}^2$, donde se encuentra I^2 (es decir, las coordenadas de la forma $T^*\omega$). Obtengamos la forma inducida $T^*\omega$.

Se tiene

$$\begin{aligned}
T^*\omega &= T^*F_1T^*(dx_2)T^*(dx_3) - T^*F_2T^*(dx_1)T^*(dx_3) + T^*F_3T^*(dx_1)T^*(dx_2) \\
&= F_1 \circ T\left(\frac{\partial T_2}{\partial u}du + \frac{\partial T_2}{\partial v}dv\right)\left(\frac{\partial T_3}{\partial u}du + \frac{\partial T_3}{\partial v}dv\right) \\
&\quad - F_2 \circ T\left(\frac{\partial T_1}{\partial u}du + \frac{\partial T_1}{\partial v}dv\right)\left(\frac{\partial T_3}{\partial u}du + \frac{\partial T_3}{\partial v}dv\right) \\
&\quad + F_3 \circ T\left(\frac{\partial T_1}{\partial u}du + \frac{\partial T_1}{\partial v}dv\right)\left(\frac{\partial T_2}{\partial u}du + \frac{\partial T_2}{\partial v}dv\right) \\
&= F_1 \circ T\left(\frac{\partial T_2}{\partial u}\frac{\partial T_3}{\partial v} - \frac{\partial T_2}{\partial v}\frac{\partial T_3}{\partial u}\right)dudv \\
&\quad - F_2 \circ T\left(\frac{\partial T_1}{\partial u}\frac{\partial T_3}{\partial v} - \frac{\partial T_1}{\partial v}\frac{\partial T_3}{\partial u}\right)dudv \\
&\quad + F_3 \circ T\left(\frac{\partial T_1}{\partial u}\frac{\partial T_2}{\partial v} - \frac{\partial T_1}{\partial v}\frac{\partial T_2}{\partial u}\right)dudv \\
&= F_1 \circ T\frac{\partial(T_2, T_3)}{\partial(u, v)}dudv + F_2 \circ T\frac{\partial(T_3, T_1)}{\partial(u, v)}dudv + F_3 \circ T\frac{\partial(T_1, T_2)}{\partial(u, v)}dudv \\
&= F \circ T \cdot \left(\frac{\partial T}{\partial u} \times \frac{\partial T}{\partial v}\right)dudv
\end{aligned}$$

Entonces la integral de la 2-forma ω sobre el 2-cubo singular T es

$$\int_T \omega = \int_{I^2} T^*\omega = \int_{I^2} F \circ T \cdot \left(\frac{\partial T}{\partial u} \times \frac{\partial T}{\partial v}\right)du\, dv$$

que no es más que la definición de integral de superficie del campo F sobre la superficie $T\colon [0, 1] \times [0, 1] \to \mathbb{R}^3$. ■

Los dos ejemplos anteriores muestran entonces que la definición de integral de una p-forma ω en un abierto $U \subseteq \mathbb{R}^n$ sobre un p-cubo singular en U, es una generalización de los conceptos de integrales de línea y de superficie de campos vectoriales ya estudiados en los capítulos 7 y 9, respectivamente. Esta es justamente la idea que perseguíamos en este capítulo: tener una teoría globalizadora de los conceptos de integral previamente estudiados.

Veamos algunos ejemplos adicionales.

Ejemplo 3. Sea ω la 3-forma en $\mathbb{R}^4$ dada por

$$\omega = x_1x_3dx_1dx_2dx_3 + (x_2 - x_4)dx_1dx_3dx_4$$

y sea $T\colon I^3 \to \mathbb{R}^4$ el 3-cubo singular dado por

$$T(y_1, y_2, y_3) = (y_3, y_2 - y_3, y_3 + y_1, y_1) = (x_1, x_2, x_3, x_4)$$

Calculemos la integral de la 3-forma ω sobre el 3-cubo singular T. Se tiene

$$\begin{aligned} T^*\omega &= T^*(x_1x_3)T^*(dx_1)T^*(dx_2)T^*(dx_3) + T^*(x_2 - x_4)T^*(dx_1)T^*(dx_3)T^*(dx_4) \\ &= y_3(y_3 + y_1)(dy_3)(dy_2 - dy_3)(dy_3 + dy_1) + (y_2 - y_3 - y_1)(dy_3)(dy_3 + dy_1)(dy_1) \\ &= y_3(y_3 + y_1)\, dy_3 dy_2 dy_1 = -y_3(y_3 y_1)\, dy_1 dy_2 dy_3 \end{aligned}$$

Entonces

$$\int_T \omega = \int_{I^3} T^*\omega = \int_0^1\int_0^1\int_0^1 -y_3(y_3y_1)\, dy_1dy_2dy_3 = -7/12 \qquad \blacksquare$$

Ejemplo 4. Sea ω la 2-forma en $\mathbb{R}^5$ dada por

$$\omega = x_3x_4x_5dx_2dx_3 + (x_1x_2 + x_4)dx_3dx_5 + x_5dx_1dx_4$$

y sea $T\colon I^2 \to \mathbb{R}^5$ el 2-cubo singular en $\mathbb{R}^5$

$$T(y_1, y_2) = (y_2, y_1, y_1y_2, y_1^2, y_2^2) = (x_1, x_2, x_3, x_4, x_5)$$

Calculemos la integral de ω sobre T. Se tiene

$$\begin{aligned} T^*\omega &= T^*(x_3x_4x_5)T^*(dx_2)T^*(dx_3) + T^*(x_1x_2 + x_4)T^*(dx_3)T^*(dx_5) + T^*(x_5)T^*(dx_1)T^*(dx_4) \\ &= (y_1y_2)(y_1^2)(y_2^2)(dy_1)(y_1dy_2 + y_2dy_1) + (y_2y_1 + y_1^2)(y_1dy_2 + y_2dy_1)(2y_2dy_2) \\ &\quad + (y_2^2)(dy_2)(2y_1dy_1) \\ &= y_1^4y_2^3dy_1dy_2 + 2y_2^2(y_2y_1 + y_1^2)dy_1dy_2 + 2y_1y_2^2dy_2dy_1 \\ &= \left(y_1^4y_2^3 + 2y_2^2(y_2y_1 + y_1^2) - 2y_1y_2^2\right)dy_1dy_2 \end{aligned}$$

Entonces

$$\int_T \omega = \int_{I^2} T^*\omega = \int_0^1\int_0^1 (y_1^4y_2^3 + 2y_2^2(y_2y_1 + y_1^2) - 2y_1y_2^2)dy_1dy_2 = 17/90 \qquad \blacksquare$$

Ejercicios (Capítulo 10, Sección 4)

1. Sean ω_1 y ω_2 dos p-formas definidas en el conjunto abierto $U \subseteq \mathbb{R}^n$, y sea $T\colon I^p \to U$, un p-cubo singular. Demuestre que

$$\int_T \omega_1 + c\omega_2 = \int_T \omega_1 + c\int_T \omega_2$$

donde $c \in \mathbb{R}$.

En los ejercicios 2–10, halle la integral de la p-forma dada, sobre el p-cubo singular indicado.

2. $\omega = dx_1 + dx_2$, $T: I^1 \to \mathbb{R}^2$, $T(y) = (y, y^2)$.

3. $\omega = x_1x_2dx_1 + x_2dx_2$, $T: I^1 \to \mathbb{R}^2$, $T(y) = (2y, 3y)$.

4. $\omega = x_1dx_1 + x_2dx_2 + x_3dx_3$, $T: I^1 \to \mathbb{R}^3$, $T(y) = (y, y^2, y^3)$.

5. $\omega = x_1x_2dx_1dx_2$, $T: I^2 \to \mathbb{R}^2$, $T(y_1, y_2) = (y_1 + y_2, y_1 - y_2)$.

6. $\omega = x_1dx_1dx_2 + x_2^2dx_1dx_3$, $T: I^2 \to \mathbb{R}^3$, $T(y_1, y_2) = (y_1, y_1y_2, y_2)$.

7. $\omega = (3x_1 + x_2)dx_1dx_3$, $T: I^2 \to \mathbb{R}^3$, $T(y_1, y_2) = (y_1 + y_2, y_2, y_1)$.

8. $\omega = x_1x_2x_3dx_1dx_2dx_3$, $T: I^3 \to \mathbb{R}^3$, $T(y_1, y_2, y_3) = (y_1y_2y_3, y_2y_3, y_3)$.

9. $\omega = x_1^2dx_1dx_2dx_4 + x_2x_4dx_2dx_3dx_4$, $T: I^3 \to \mathbb{R}^4$, $T(y_1, y_2, y_3) = (y_1y_2y_3, y_2y_3, y_1 + y_2, y_1y_2)$.

10. $\omega = dx_1dx_2dx_3dx_5 - x_1x_2x_5dx_2dx_3dx_4dx_5$, $T: I^4 \to \mathbb{R}^5$, $T(y_1, y_2, y_3, y_4) = (y_4, y_2y_3, y_1y_4, y_4^2, y_4^3)$.

10.5 Integración de p-formas sobre p-cadenas

En esta sección estudiaremos integrales de p-formas ω definidas en conjuntos abiertos U de $\mathbb{R}^n$, sobre regiones más generales que los p-cubos singulares considerados en la sección anterior. Estas regiones se llaman "p-cadenas singulares", las cuales se presentan a continuación.

Sea R un subconjunto de $\mathbb{R}^n$. Una *p-cadena singular en R* (o simplemente p-cadena en R), que denotamos por c, es una expresión del tipo

$$c = a_1T_1 + a_2T_2 + \cdots + a_kT_k$$

donde a_i son números reales dados, y $T_i: I^p \to R$ son p-cubos singulares en R, $i = 1, 2, \ldots, k$. Así pues, una p-cadena singular es una "combinación lineal" de p-cubos singulares (ver nota después de la definición siguiente).

Si ω es una p-forma definida en el conjunto abierto $U \subseteq \mathbb{R}^n$ y $c = \sum_{i=1}^{k} a_iT_i$ es una p-cadena singular en U, definimos la *integral de ω sobre c*, denotada por $\int_c \omega$, como

$$\int_c \omega = \sum_{i=1}^{k} a_i \int_{T_i} \omega$$

NOTA. Una p-cadena singular se ha presentado como "una combinación lineal" de p-cubos singulares, cada uno de los cuales es una función del tipo $T_i: I^p \to \mathbb{R}^n$. NO SE DEBE INTERPRETAR UNA P-CADENA SINGULAR COMO UNA FUNCION (ver ejercicio 1 al final de la sección). O sea, no es correcto decir, por ejemplo, que la p-cadena singular $c = 3T_1 - 5T_2$, donde T_1 y T_2 son p-cubos singulares, es tal que $c(x) = 3T_1(x) - 5T_2(x)$, $x \in I^p$. Debemos tomar a las p-cadenas singulares como *expresiones formales*, las cuales "parecen" combinaciones lineales de p-cubos singulares. El hecho es que estos objetos formales se usarán como regiones de integración de formas diferenciales. La definición de integral de una forma sobre una p-cadena singular ya se dio anteriormente, y está escrita en términos de integrales de la forma sobre los p-cubos que forman la p-cadena. Estas últimas integrales ya quedaron definidas en la sección anterior.

Veamos algunos ejemplos.

Ejemplo 1. Sea ω la 2-forma en $\mathbb{R}^3$ dada por $\omega = x_1x_3dx_1dx_2 + x_2dx_1dx_3$ y sea la 2-cadena singular en $\mathbb{R}^3$, $c = 3T_1 + 2T_2$, en donde T_1 y T_2 son los 2-cubos singulares en $\mathbb{R}^3$ dados por $T_1\colon I^2 \to \mathbb{R}^3$, $T_1(y_1, y_2) = (y_2, y_1, y_1y_2)$, $T_2\colon I^2 \to \mathbb{R}^3$, $T_2(y_1, y_2) = (y_1^2, y_1y_2^2, y_1^2y_2)$. Calculemos la integral de la 2-forma ω sobre la 2-cadena c. Se tiene

$$\int_c \omega = 3\int_{T_1} \omega + 2\int_{T_2} \omega = 3\int_{I^2} T_1^*\omega + 2\int_{I^2} T_2^*\omega$$

Calculemos $T_1^*\omega$ y $T_2^*\omega$. Se tiene

$$\begin{aligned} T_1^*\omega &= T_1^*(x_1x_3)T_1^*(dx_1)T_1^*(dx_2) + T_1^*(x_2)T_1^*(dx_1)T_1^*(dx_3) \\ &= (y_2)(y_1y_2)(dy_2)(dy_1) + (y_1)(dy_2)(y_1dy_2 + y_2dy_1) = -(y_1y_2^2 + y_1y_2)dy_1dy_2 \end{aligned}$$

y

$$\begin{aligned} T_2^*\omega &= T_2^*(x_1x_3)T_2^*(dx_1)T_2^*(dx_2) + T_2^*(x_2)T_2^*(dx_1)T_2^*(dx_3) \\ &= (y_1^2)(y_1^2y_2)(2y_1dy_1)(y_2^2dy_1 + 2y_1y_2dy_2) + (y_1y_2^2)(2y_1dy_1)(2y_1y_2dy_1 + y_1^2dy_2) \\ &= (4y_1^6y_2^2 + 2y_1^4y_2^2)dy_1dy_2 \end{aligned}$$

Entonces

$$\begin{aligned} \int_c \omega = 3\int_{I^2} T_1^*\omega + 2\int_{I^2} T_2^*\omega &= 3\int_0^1\int_0^1 -(y_1y_2^2 + y_1y_2)dy_1dy_2 + 2\int_0^1\int_0^1 (4y_1^6y_2^2 + 2y_1^4y_2^2)dy_1dy_2 \\ &= 3(-5/12) + 2(34/105) = -253/420 \end{aligned}$$

∎

Las p-cadenas singulares se pueden operar en forma algebraica entre sí de la manera natural: si $c = a_1T_1 + a_2T_2 + \cdots + a_kT_k$, y $d = b_1T_1 + b_2T_2 + \cdots + b_mT_m$ son dos p-cadenas singulares en U, entonces su suma es la p-cadena singular

$$c + d = a_1T_1 + \cdots + a_kT_k + b_1T_1 + \cdots + b_mT_m$$

y el producto de la p-cadena singular c por el número real α es la p-cadena singular

$$\alpha c = \alpha a_1T_1 + \alpha a_2T_2 + \cdots + \alpha a_kT_k$$

En el teorema siguiente se recogen las principales propiedades de las integrales de p-formas sobre p-cadenas singulares.

Teorema 10.5.1 Sean ω_1 y ω_2 dos p-formas definidas en el conjunto abierto U de $\mathbb{R}^n$, sean c y d dos p-cadenas singulares en U, y α un número real. Entonces:

a. $\displaystyle\int_c (\omega_1 + \omega_2) = \int_c \omega_1 + \int_c \omega_2$

b. $\displaystyle\int_{c+d} \omega_1 = \int_c \omega_1 + \int_d \omega_1$

c. $\displaystyle\int_{\alpha c} \omega_1 = \alpha\int_c \omega_1 = \int_c \alpha\omega_1$

Demostración. Se trata de verificaciones de simple rutina, que se obtienen directamente de las definiciones correspondientes y de las propiedades previamente establecidas para la operación estrella en formas. Presentamos las demostraciones de las propiedades de los incisos **a.** y **b.** y dejamos para el lector la del inciso **c.** Sea $c = a_1T_1 + a_2T_2 + \cdots + a_kT_k$. Se tiene

$$\begin{aligned}\int_c(\omega_1+\omega_2) &= \sum_{i=1}^{k} a_i \int_{T_i}(\omega_1+\omega_2) = \sum_{i=1}^{k} a_i \int_{I^p} T_i^*(\omega_1+\omega_2)\\ &= \sum_{i=1}^{k} a_i \int_{I^p}(T_i^*\omega_1 + T_i^*\omega_2) = \sum_{i=1}^{k} a_i\left(\int_{I^p} T_i^*\omega_1 + \int_{I^p} T_i^*\omega_2\right)\\ &= \sum_{i=1}^{k} a_i \int_{I^p} T_i^*\omega_1 + \sum_{i=1}^{k} a_i \int_{I^p} T_i^*\omega_2 = \sum_{i=1}^{k} a_i \int_{T_i}\omega_1 + \sum_{i=1}^{k} a_i \int_{T_i}\omega_2\\ &= \int_c \omega_1 + \int_c \omega_2\end{aligned}$$

(donde usamos la propiedad de linealidad de las integrales p-múltiples estudiadas en el capítulo 6). Si $d = b_1T_1 + \cdots + b_mT_m$, se tiene $c + d = a_1T_1 + \cdots + a_kT_k + b_1T_1 + \cdots + b_mT_m$, y entonces

$$\int_{c+d}\omega_1 = \sum_{i=1}^{k} a_i \int_{T_i}\omega_1 + \sum_{j=1}^{m} b_j \int_{T_j}\omega_1 = \int_c \omega_1 + \int_d \omega_1 \qquad \text{Q.E.D}$$

Estudiaremos ahora un concepto que tiene que ver de alguna manera con "la frontera" de una p-cadena. Sea $p \in \mathbb{N}$. Para cada i, $1 \le i \le p$, definimos las funciones $\partial_i, \partial_i'\colon I^{p-1} \to I^p$, como

$$\begin{aligned}\partial_i(t_1, t_2, \ldots, t_{p-1}) &= (t_1, \ldots, t_{i-1}, 0, t_i, \ldots, t_{p-1})\\ \partial_i'(t_1, t_2, \ldots, t_{p-1}) &= (t_1, \ldots, t_{i-1}, 1, t_i, \ldots, t_{p-1})\end{aligned}$$

Veamos un par de ejemplos que nos muestran "qué es lo que hacen" estas funciones

Ejemplo 2. Sea $p = 2$. Tenemos entonces 4 funciones de las definidas anteriormente, ∂_1, ∂_2, ∂_1', $\partial_2'\colon [0, 1] \to [0, 1] \times [0, 1]$, a saber $\partial_1(t) = (0, t)$, $\partial_2(t) = (t, 0)$, $\partial_1'(t) = (1, t)$, $\partial_2'(t) = (t, 1)$. Obsérvese entonces que cada una de estas funciones se encarga de "incluír" o de "copiar" al 1-cubo $[0, 1]$ en cada una de los 4 lados del 2-cubo $[0, 1] \times [0, 1]$, a saber: ∂_1 copia el $[0, 1]$ en el lado izquierdo del 2-cubo $[0, 1] \times [0, 1]$ (además, esta copia se hace iniciando en $(0, 0)$ y terminando en $(0, 1)$); ∂_2 lo copia en el lado inferior (del $(0, 0)$ al $(1, 0)$); ∂_1' lo copia en el lado derecho (del $(1, 0)$ al $(1, 1)$) y ∂_2' lo copia en el lado superior (del $(0, 1)$ al $(1, 1)$). ■

Ejemplo 3. Sea $p = 3$. Se tienen 6 funciones del 2-cubo $I^2 = [0, 1] \times [0, 1]$ en el 3-cubo $I^3 = [0, 1] \times [0, 1] \times [0, 1]$. Estas son $\partial_1(t_1, t_2) = (0, t_1, t_2)$, $\partial_2(t_1, t_2) = (t_1, 0, t_2)$, $\partial_3(t_1, t_2) = (t_1, t_2, 0)$, $\partial_1'(t_1, t_2) = (1, t_1, t_2)$, $\partial_2'(t_1, t_2) = (t_1, 1, t_2)$, $\partial_3'(t_1, t_2) = (t_1, t_2, 1)$. Nótese de nuevo que estas funciones "copian" el 2-cubo I^2 en cada una de las 6 caras del 3-cubo I^3. Por ejemplo ∂_3 se lleva el 2-cubo I^2 "al piso" del 3 cubo I^3, y ∂_3' se lleva a I^2 "al techo" de I^3. ■

Los ejemplos anteriores nos muestran, en situaciones concretas, que las funciones ∂_i, ∂_i': $I^{p-1} \to I^p$ definidas arriba son simplemente "inclusiones" o "copias" del $(p-1)$-cubo I^{p-1} en las diferentes "caras" del p-cubo I^p.

Nótese que si $T: I^p \to \mathbb{R}^n$ es un p-cubo singular en $\mathbb{R}^n$, y $\partial_i: I^{p-1} \to I^p$ es una de las funciones ya consideradas anteriormente, podemos formar la composición $T \circ \partial_i: I^{p-1} \to \mathbb{R}^n$, y ésta será un $(p-1)$-cubo singular en $\mathbb{R}^n$, el cual es esencialmente el mismo p-cubo T restringido a "las caras" de I^p (que son $(p-1)$-cubos). Estamos ahora en posibilidades de definir un concepto muy importante para p-cubos singulares, a saber, su frontera.

Sea $T: I^p \to U$ un p-cubo singular en el conjunto abierto $U \subseteq \mathbb{R}^n$. Definimos la *frontera de* T, denotada por ∂T, como la $(p-1)$-cadena singular

$$\partial T = \sum_{i=1}^{p} (-1)^i (T \circ \partial_i - T \circ \partial_i')$$

La idea de esta definición es restringir el p-cubo T a las caras de I^p, quedando así una combinación lineal de $(p-1)$-cubos (a saber, las composiciónes de T con las funciones ∂_i y ∂_i'), las cuales forman una $(p-1)$-cadena. La introducción de los signos menos en esta definición es para asegurar que el recorrido quede "orientado".

Veamos funcionar estas ideas con un ejemplo concreto.

Ejemplo 4. Sea $p = 2$. La frontera de un 2-cubo $T: I^2 \to \mathbb{R}^n$ es

$$\partial T = \sum_{i=1}^{2} (-1)^i (T \circ \partial_i - T \circ \partial_i') = -T \circ \partial_1 + T \circ \partial_1' + T \circ \partial_2 - T \circ \partial_2'$$

La intención de los signos menos en el primer y último sumandos es *invertir* el recorrido de (las restricciones de T a) los lados de I^2, de modo que éste se realice de tal manera que el interior de I^2 quede siempre a su izquierda (es decir, de modo que la frontera de I^2 quede orientada positivamente). Dejamos al lector que verifique detenidamente este hecho, comparando con los sentidos de recorrido mencionados en el ejemplo 2. ■

En general, para una p-cadena $c = a_1 T_1 + \cdots + a_k T_k$, se define la *frontera de* c, denotada por ∂c, como

$$\partial c = a_1 \partial T_1 + \cdots + a_k \partial T_k$$

Ejemplo 5. Partamos de un 3-cubo $T: I^3 \to \mathbb{R}^n$. Su frontera es la 2-cadena

$$\begin{aligned}\partial T &= \sum_{i=1}^{3} (-1)^i (T \circ \partial_i - T \circ \partial_i') \\ &= -T \circ \partial_1 + T \circ \partial_1' + T \circ \partial_2 - T \circ \partial_2' - T \circ \partial_3 + T \circ \partial_3'\end{aligned}$$

donde ∂_1, ∂_1', ∂_2, ∂_2', ∂_3, ∂_3': $I^2 \to I^3$ son las funciones

$$\begin{aligned}\partial_1(t_1, t_2) &= (0, t_1, t_2) \qquad & \partial_1'(t_1, t_2) &= (1, t_1, t_2) \\ \partial_2(t_1, t_2) &= (t_1, 0, t_2) \qquad & \partial_2'(t_1, t_2) &= (t_1, 1, t_2) \\ \partial_3(t_1, t_2) &= (t_1, t_2, 0) \qquad & \partial_3'(t_1, t_2) &= (t_1, t_2, 1)\end{aligned}$$

(ver ejemplo 3). La frontera de la 2-cadena ∂T es

$$\partial\partial T = -\partial(T \circ \partial_1) + \partial(T \circ \partial_1') + \partial(T \circ \partial_2) - \partial(T \circ \partial_2') - \partial(T \circ \partial_3) + \partial(T \circ \partial_3')$$

Nótese que cada uno de los sumandos de la expresión anterior se refiere a la frontera de un 2-cubo, por ejemplo, $T \circ \partial_1 \colon I^2 \to \mathbb{R}^n$. Llamemos $\tilde{\partial}_1, \tilde{\partial}_2, \tilde{\partial}_1', \tilde{\partial}_2' \colon I^1 \to I^2$, a las funciones $\tilde{\partial}_1(t) = (0, t)$, $\tilde{\partial}_2(t) = (t, 0)$, $\tilde{\partial}_1'(t) = (1, t)$, $\tilde{\partial}_2'(t) = (t, 1)$ (ver ejemplo 2). Entonces, se tiene que (el primer sumando de $\partial\partial T$ es)

$$\begin{aligned}
\partial(T \circ \partial_1) &= \sum_{i=1}^{2}(-1)^i((T \circ \partial_1) \circ \tilde{\partial}_i - (T \circ \partial_1) \circ \tilde{\partial}_i') \\
&= -(T \circ \partial_1) \circ \tilde{\partial}_1 + (T \circ \partial_1) \circ \tilde{\partial}_1' + (T \circ \partial_1) \circ \tilde{\partial}_2 - (T \circ \partial_1) \circ \tilde{\partial}_2' \\
&= -T \circ (\partial_1 \circ \tilde{\partial}_1) + T \circ (\partial_1 \circ \tilde{\partial}_1') + T \circ (\partial_1 \circ \tilde{\partial}_2) - T \circ (\partial_1 \circ \tilde{\partial}_2') \\
&= -T(0, 0, t) + T(0, 1, t) + T(0, t, 0) - T(0, t, 1)
\end{aligned}$$

Haciendo operaciones similares se llega a que (los siguientes sumandos de $\partial\partial T$ son)

$$\begin{aligned}
\partial(T \circ \partial_1') &= -T(1, 0, t) + T(1, 1, t) + T(1, t, 0) - T(1, t, 1) \\
\partial(T \circ \partial_2) &= -T(0, 0, t) + T(1, 0, t) + T(t, 0, 0) - T(t, 0, 1) \\
\partial(T \circ \partial_2') &= -T(0, 1, t) + T(1, 1, t) + T(t, 1, 0) - T(t, 1, 1) \\
\partial(T \circ \partial_3) &= -T(0, t, 0) + T(1, t, 0) + T(t, 0, 0) - T(t, 1, 0) \\
\partial(T \circ \partial_3') &= -T(0, t, 1) + T(1, t, 1) + T(t, 0, 1) - T(t, 1, 1)
\end{aligned}$$

Entonces

$$\begin{aligned}
\partial\partial T &= -\partial(T \circ \partial_1) + \partial(T \circ \partial_1') + \partial(T \circ \partial_2) - \partial(T \circ \partial_2') - \partial(T \circ \partial_3) + \partial(T \circ \partial_3') \\
&= T(0, 0, t) - T(0, 1, t) - T(0, t, 0) + T(0, t, 1) \\
&\quad - T(1, 0, t) + T(1, 1, t) + T(1, t, 0) - T(1, t, 1) \\
&\quad - T(0, 0, t) + T(1, 0, t) + T(t, 0, 0) - T(t, 0, 1) \\
&\quad + T(0, 1, t) - T(1, 1, t) - T(t, 1, 0) + T(t, 1, 1) \\
&\quad + T(0, t, 0) - T(1, t, 0) - T(t, 0, 0) + T(t, 1, 0) \\
&\quad - T(0, t, 1) + T(1, t, 1) + T(t, 0, 1) - T(t, 1, 1) \\
&= 0
\end{aligned}$$

■

La situación mostrada en el ejemplo anterior es un hecho general que se incluye en el teorema siguiente.

Teorema 10.5.2 Sean c y d p-cadenas singulares en el conjunto abierto U de $\mathbb{R}^n$, y sea α un número real. Entonces

a. $\partial(\alpha c + d) = \alpha\partial c + \partial d$

b. $\partial\partial c = 0$

Demostración. La demostración del inciso **a** es directa: sea $c = a_1 T_1 + \cdots + a_k T_k$ y $d = b_1 T_1 + \cdots + b_m T_m$. Entonces

$$\begin{aligned}\partial(\alpha c + d) &= \alpha a_1 \partial T_1 + \cdots + \alpha a_k \partial T_k + b_1 \partial T_1 + \cdots + b_m \partial T_m \\ &= \alpha\big(a_1 \partial T_1 + \cdots + a_k \partial T_k\big) + b_1 \partial T_1 + \cdots + b_m \partial T_m \\ &= \alpha \partial c + \partial d\end{aligned}$$

La demostración del inciso **b** tiene algunas complicaciones respecto al manejo de expresiones con muchos sumandos y muchos índices (el lector lo puede llegar a presentir del caso -muy sencillo- presentado en el ejemplo 5). Quede entonces para el lector a quien le gusten los retos operativos.
Q.E.D.

Establezcamos un resultado adicional que se usará en la demostración del teorema de la siguiente sección.

Teorema 10.5.3 Sean $\partial_i, \partial_i' \colon I^{p-1} \to I^p$, las funciones (definidas anteriormente)

$$\begin{aligned}\partial_i(t_1, t_2, \ldots, t_{p-1}) &= (t_1, \ldots, t_{i-1}, 0, t_i, \ldots, t_{p-1}) \\ \partial_i'(t_1, t_2, \ldots, t_{p-1}) &= (t_1, \ldots, t_{i-1}, 1, t_i, \ldots, t_{p-1})\end{aligned}$$

Entonces

$$(\partial_i^*)(dx_j) = (\partial_i')^*(dx_j) = \begin{cases} dx_j & \text{si } j < i \\ 0 & \text{si } j = i \\ dx_{j-1} & \text{si } j > i \end{cases}$$

Demostración. Con nuestra notación clásica de la sección 3, denotando con y's a las coordenadas de I^{p-1} (el dominio de ∂_i y ∂_i') y con x's a las coordenadas de I^p, escribiríamos

$$\begin{aligned}\partial_i(y_1, y_2, \ldots, y_{p-1}) &= (y_1, \ldots, y_{i-1}, 0, y_i, \ldots, y_{p-1}) = (x_1, \ldots, x_{i-1}, \ldots, x_p) \\ \partial_i'(y_1, y_2, \ldots, y_{p-1}) &= (y_1, \ldots, y_{i-1}, 1, y_i, \ldots, y_{p-1}) = (x_1, \ldots, x_{i-1}, \ldots, x_p)\end{aligned}$$

La j-ésima función coordenada de ∂_i es entonces $x_j = y_j$ si $j < i$, $x_j = 0$ si $j = i$, $x_j = y_{j+1}$ si $j > i$. En el caso de ∂_i', sus funciones coordenadas son $x_j = y_j$ si $j < i$, $x_j = 1$ si $j = i$, $x_j = y_{j-1}$ si $j > i$. En ambos casos se tiene entonces que si $j < i$, $(\partial_i^*)(dx_j) = (\partial_i'^*)(dx_j) = dy_j$, si $j = i (\partial_i^*)(dx_j) = (\partial_i'^*)(dx_j) = 0$, y si $j > i$, $(\partial_i^*)(dx_j) = (\partial_i'^*)(dx_j) = dy_{j-1}$. Esto es lo que se quería probar.
Q.E.D.

A manera de preámbulo de la siguiente (y última) sección, veamos el ejemplo siguiente.

Ejemplo 6. Considere la 1-forma ω en $\mathbb{R}^3$ dada por

$$\omega = x_1 x_2 dx_1 + x_1^2 x_2 x_3 dx_2 + x_1 x_3^2 dx_3$$

Sea c la 2-cadena singular

$$c = 3T_1 - 7T_2$$

donde $T_1, T_2: I^2 \to \mathbb{R}^3$ son los 2-cubos singulares

$$T_1(y_1, y_2) = (y_1 y_2, y_2^2, y_1), \qquad T_2(y_1, y_2) = (y_2, y_1, y_1 y_2)$$

Queremos calcular la integral de la 1-forma ω sobre la frontera de la 2-cadena c (la cual es una 1-cadena). La frontera de c es $\partial c = 3\partial T_1 - 7\partial T_2$, donde

$$\partial T_1 = \sum_{i=1}^{2}(-1)^i(T_1 \circ \partial_i - T_1 \circ \partial_i') = -T_1 \circ \partial_1 + T_1 \circ \partial_1' + T_1 \circ \partial_2 - T_1 \circ \partial_2'$$

$$\partial T_2 = \sum_{i=1}^{2}(-1)^i(T_2 \circ \partial_i - T_2 \circ \partial_i') = -T_2 \circ \partial_1 + T_2 \circ \partial_1' + T_2 \circ \partial_2 - T_2 \circ \partial_2'$$

siendo $\partial_1, \partial_2, \partial_1', \partial_2': I^1 \to I^2$ las funciones $\partial_1(t) = (0, t)$, $\partial_2(t) = (t, 0)$, $\partial_1'(t) = (1, t)$, $\partial_2'(t) = (t, 1)$ (ver ejemplo 2). Entonces

$$\partial c = 3\partial T_1 - 7\partial T_2 = \sum_{i=1}^{2}(-1)^i\big(3(T_1 \circ \partial_i - T_1 \circ \partial_i') - 7(T_2 \circ \partial_i - T_2 \circ \partial_i')\big)$$

La integral de ω sobre ∂c es entonces

$$\begin{aligned}
\int_{\partial c} \omega &= \sum_{i=1}^{2}(-1)^i \int_0^1 \big(3(T_1 \circ \partial_i - T_1 \circ \partial_i') - 7(T_2 \circ \partial_i - T_2 \circ \partial_i')\big)^* \omega \\
&= \sum_{i=1}^{2}(-1)^i \int_0^1 \big(3(T_1 \circ \partial_i - T_1 \circ \partial_i')^* - 7(T_2 \circ \partial_i - T_2 \circ \partial_i')\big)^* \omega \\
&= \sum_{i=1}^{2}(-1)^i \int_0^1 \big(3((T_1 \circ \partial_i)^* - (T_1 \circ \partial_i')^*) - 7((T_2 \circ \partial_i)^* - (T_2 \circ \partial_i')^*)\big) \omega \\
&= \sum_{i=1}^{2}(-1)^i \int_0^1 \big(3(\partial_i^* \circ T_1^* \omega - \partial_i'^* \circ T_1^* \omega) - 7(\partial_i^* \circ T_2^* \omega - \partial_i'^* \circ T_2^* \omega)\big)
\end{aligned}$$

Obtengamos las formas inducidas $T_1^*\omega$ y $T_2^*\omega$. Se tiene

$$\begin{aligned}
T_1^*\omega &= T_1^*(x_1x_2)T_1^*(dx_1) + T_1^*(x_1^2x_2x_3)T_1^*(dx_2) + T_1^*(x_1x_3^2)T_1^*(dx_3) \\
&= (y_1y_2)(y_2^2)(y_1dy_2 + y_2dy_1) + (y_1y_2)^2(y_2^2)(y_1)(2y_2dy_2) + (y_1y_2)(y_1^2)(dy_1) \\
&= (y_1y_2^4 + y_1^3y_2)dy_1 + (y_1^2y_2^3 + 2y_1^3y_2^5)dy_2
\end{aligned}$$

y

$$\begin{aligned}
T_2^*\omega &= T_2^*(x_1x_2)T_2^*(dx_1) + T_2^*(x_1^2x_2x_3)T_2^*(dx_2) + T_2^*(x_1x_3^2)T_2^*(dx_3) \\
&= (y_2)(y_1)(dy_2) + (y_2^2)(y_1)(y_1y_2)(dy_1) + (y_2)(y_1y_2)^2(y_1dy_2 + y_2dy_1) \\
&= (y_1^2y_2^3 + y_1^2y_2^4)dy_1 + (y_1y_2 + y_1^3y_2^3)dy_2
\end{aligned}$$

Entonces

$$\partial_1^* \circ T_1^* \omega = \partial_1^*(T_1^* \omega) = \partial_1^*(y_1 y_2^4 + y_1^3 y_2)\partial_1^*(dy_1) + \partial_1^*(y_1^2 y_2^3 + 2y_1^3 y_2^5)\partial_1^*(dy_2) = 0$$

Procediendo de esta manera se obtiene que

$$\begin{aligned}
\partial_2^* \circ T_1^* \omega &= \partial_2^*(T_1^* \omega) = 0 \\
\partial_1'^* \circ T_1^* \omega &= \partial_1'^*(T_1^* \omega) = (t^3 + 2t^5)dt \\
\partial_2'^* \circ T_1^* \omega &= \partial_2'^*(T_1^* \omega) = (t + t^3)dt \\
\partial_1^* \circ T_2^* \omega &= \partial_1^*(T_2^* \omega) = 0 \\
\partial_2^* \circ T_2^* \omega &= \partial_2^*(T_2^* \omega) = 0 \\
\partial_1'^* \circ T_2^* \omega &= \partial_1'^*(T_2^* \omega) = (t + t^3)dt \\
\partial_2'^* \circ T_2^* \omega &= \partial_2'^*(T_2^* \omega) = 2t^2 dt
\end{aligned}$$

Entonces

$$\begin{aligned}
\int_{\partial c} \omega &= \sum_{i=1}^{2}(-1)^i \int_0^1 \left(3(\partial_i^* \circ T_1^* \omega - \partial_i'^* \circ T_1^* \omega) - 7(\partial_i^* \circ T_2^* \omega - \partial_i'^* \circ T_2^* \omega)\right) \\
&= -\int_0^1 \left(3(\partial_1^* \circ T_1^* \omega - \partial_1'^* \circ T_1^* \omega) - 7(\partial_1^* \circ T_2^* \omega - \partial_1'^* \circ T_2^* \omega)\right) \\
&\quad + \int_0^1 \left(3(\partial_2^* \circ T_1^* \omega - \partial_2'^* \circ T_1^* \omega) - 7(\partial_2^* \circ T_2^* \omega - \partial_2'^* \circ T_2^* \omega)\right) \\
&= -\int_0^1 \left(-3(t^3 + 2t^5) + 7(t + t^3)\right) dt \\
&= \int_0^1 \left(-3(t + t^3)dt + 7(2t^2)\right) dt \\
&= -13/12
\end{aligned}$$

Queremos calcular ahora la integral de la 2-forma $d\omega$ (la diferencial exterior de la 1-forma ω) sobre la 2-cadena c. Obtengamos $d\omega$.

$$\begin{aligned}
d\omega &= d(x_1 x_2)dx_1 + d(x_1^2 x_2 x_3)dx_2 + d(x_1 x_3^2)dx_3 \\
&= (2x_1 x_2 x_3 - x_1)dx_1 dx_2 + x_3^2 dx_1 dx_3 - x_1^2 x_2 dx_2 dx_3
\end{aligned}$$

Se tiene

$$\int_c d\omega = 3\int_{T_1} d\omega - 7\int_{T_2} d\omega = 3\int_{I^2} T_1^*(d\omega) - 7\int_{I^2} T_2^*(d\omega)$$

Obtengamos las formas inducidas que aparecen en esta expresión.

$$\begin{aligned}
T_1^*(d\omega) &= T_1^*(2x_1x_2x_3 - x_1)T_1^*(dx_1)T_1^*(dx_2) + T_1^*(x_3^2)T_1^*(dx_1)T_1^*(dx_3)\\
&\quad - T_1^*(x_1^2x_2)T_1^*(dx_2)T_1^*(dx_3)\\
&= \big((2y_1y_2)(y_2^2)(y_1) - y_1y_2\big)(y_1dy_2 + y_2dy_1)(2y_2dy_2)\\
&\quad + (y_1^2)(y_1dy_2 + y_2dy_1)(dy_1) - (y_1^2y_2^2)(y_2^2)(2y_2dy_2)(dy_1)\\
&= (6y_1^2y_2^5 - 2y_1y_2^3 - y_1^3)dy_1dy_2
\end{aligned}$$

y

$$\begin{aligned}
T_2^*(d\omega) &= T_2^*(2x_1x_2x_3 - x_1)T_2^*(dx_1)T_2^*(dx_2) + T_2^*(x_3^2)T_2^*(dx_1)T_2^*(dx_3)\\
&\quad - T_2^*(x_1^2x_2)T_2^*(dx_2)T_2^*(dx_3)\\
&= \big(2(y_2)(y_1)(y_1y_2) - y_2\big)(dy_2)(dy_1) + (y_1y_2)^2(dy_2)(y_1dy_2 + y_2dy_1)\\
&\quad - (y_2^2)(y_1)(dy_1)(y_1dy_2 + y_2dy_1)\\
&= (-3y_1^2y_2^2 + y_2 - y_1^2y_2^3)dy_1dy_2
\end{aligned}$$

Entonces

$$\begin{aligned}
\int_c d\omega &= 3\int_{I^2} T_1^*(d\omega) - 7\int_{I^2} T_2^*(d\omega)\\
&= 3\int_0^1\int_0^1 (6y_1^2y_2^5 - 2y_1y_2^3 - y_1^3)dy_1dy_2 - 7\int_0^1\int_0^1 (-3y_1^2y_2^2 + y_2 - y_1^2y_2^3)dy_1dy_2\\
&= 3(-1/6) - 7(1/12) = -13/12
\end{aligned}$$

Luego

$$\int_{\partial c}\omega = -13/12 = \int_c d\omega$$

Esta no es una mera coincidencia. Es un hecho general: el resultado más importante de todo el libro, que se expondrá en la próxima sección. ∎

Ejercicios (Capítulo 10, Sección 5)

1. Considere la 1-forma en $\mathbb{R}^2$, $\omega = x_1dx_1 + x_1x_2dx_2$, y la 1-cadena singular en $\mathbb{R}^2$, $c = T_1 + 4T_2$, en donde T_1, T_2: $I^1 \to \mathbb{R}^2$, son los 1-cubos singulares $T_1(y) = (y^2, 3y)$, $T_2(y) = (4y, y^3)$. La integral de la forma ω sobre la 1-cadena c es, por definición

$$\int_c \omega = \int_{T_1} \omega + 4\int_{T_2} \omega = \int_{I^1} T_1^*\omega + 4\int_{I^1} T_2^*\omega$$

$$= \int_0^1 (y^2)(2y\,dy) + (y^2)(3y)(3\,dy) + 4\int_0^1 (4y)(4\,dy) + (4y)(y^3)(3y^2dy)$$

$$= \int_0^1 (11y^3 + 64y + 48y^6)\,dy = \frac{1165}{28}$$

Por otra parte, la 1-cadena singular c se puede escribir como

$$c(y) = (T_1 + 4T_2)(y) = T_1(y) + 4T_2(y) = (y^2, 3y) + 4(4y, y^3)$$
$$= (y^2 + 16y, 3y + 4y^3)$$

Ésta es entonces un nuevo 1-cubo singular en $\mathbb{R}^2$, digamos $\tilde{T}: I^1 \to \mathbb{R}^2$, $\tilde{T}(y) = (y^2 + 16y, 3y + 4y^3)$. Así, la integral de la forma ω sobre c es

$$\int_c \omega = \int_{\tilde{T}} \omega = \int_{I^1} \tilde{T}^*\omega$$

$$= \int_0^1 (y^2 + 16y)(2y + 16)\,dy + (y^2 + 16y)(3y + 4y^3)(3 + 12y^2)\,dy$$

$$= \int_0^1 (48y^7 + 768y^6 + 48y^5 + 768y^4 + 11y^3 + 192y^2 + 256y)\,dy = \frac{66089}{140}$$

que es un resultado distinto al calculado anteriormente. Explique cuál de estos procedimientos es incorrecto.

En los ejercicios 2–5, considere la 1-cadena singular $c = 3T_1 - 6T_2$, en donde $T_1, T_2: I^1 \to \mathbb{R}^3$ son los 1-cubos singulares en $\mathbb{R}^3$ dados por $T_1(y) = (y, 3y, y^2)$, $T_2(y) = (3y, 2y, y)$. Calcule la integral de la 1-forma ω dada sobre la 1-cadena c.

2. $\omega = dx_1 - dx_3$

3. $\omega = x_1x_2dx_1 + x_3dx_2$

4. $\omega = x_1dx_1 + x_2dx_2 + x_3dx_3$

5. $\omega = x_1x_2x_3dx_2$

En los ejercicios 6–10, considere la 2-cadena singular $c = T_1 - T_2 + 3T_3$, donde $T_1, T_2, T_3: I^2 \to \mathbb{R}^3$ son los 2-cubos singulares en $\mathbb{R}^3$ dados por $T_1(y_1, y_2) = (y_1, y_2, y_1y_2)$, $T_2(y_1, y_2) = (y_1 + y_2, y_1 - y_2, y_2)$, $T_3(y_1, y_2) = (3y_2, 4y_1y_2, y_1)$. Calcule la integral de la 2-forma ω dada sobre la 2-cadena singular c.

6. $\omega = dx_1dx_3$

7. $\omega = x_1dx_1dx_3 + x_3dx_2dx_3$

8. $\omega = dx_1dx_2 + x_1dx_1dx_3$

9. $\omega = x_1^2dx_1dx_2 + x_2dx_1dx_3$

10. $\omega = x_1x_2dx_1dx_2 + x_1x_3dx_1dx_3 + x_2x_3dx_2dx_3$

10.6 Teorema (general) de Stokes

Stokes' theorem shares three important attributes with many fully evolved major theorems: 1) It is trivial. 2) It is trivial because the terms appearing in it have been properly defined. 3) It has significant consequences.

Michael Spivak

Sin más preámbulos, presentamos el teorema más importante de todo el libro.

Teorema 10.6.1 (El teorema -general- de Stokes). Sea U un conjunto abierto del espacio $\mathbb{R}^n$, ω una $(p-1)$-forma en U, y c una p-cadena singular en U. Entonces

$$\int_{\partial c}\omega=\int_c d\omega$$

Demostración. Consideremos primero el caso en el que $p=n$. Entonces ω es una $(n-1)$-forma definida en un conjunto abierto de $\mathbb{R}^n$, y se escribe por lo tanto como

$$\omega=\sum_{i=1}^{n} f_i dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n$$

Tomemos la n-cadena singular $c=T$, formada por el único n-cubo singular $T\colon I^n\to\mathbb{R}^n$ dado por $T(x_1,x_2,\ldots,x_n)=(x_1,x_2,\ldots,x_n)$. Entonces

$$\begin{aligned}\int_c d\omega &= \int_c \sum_{i=1}^{n}(df_i)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n\\ &=\sum_{i=1}^{n}\int_c\left(\frac{\partial f_1}{\partial x_1}dx_1+\cdots+\frac{\partial f_i}{\partial x_i}dx_i+\cdots+\frac{\partial f_n}{\partial x_n}dx_n\right)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n\\ &=\sum_{i=1}^{n}\int_c\frac{\partial f_i}{\partial x_i}dx_i dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n\\ &=\sum_{i=1}^{n}(-1)^{i-1}\int_c\frac{\partial f_i}{\partial x_i}dx_1\cdots dx_{i-1}dx_i dx_{i+1}\cdots dx_n\\ &=\sum_{i=1}^{n}(-1)^{i-1}\int_{I^n}\frac{\partial f_i}{\partial x_i}dx_1 dx_2\cdots dx_n\end{aligned}$$

La integral que aparece en la expresión anterior es una integral n-múltiple de la función $\frac{\partial f_i}{\partial x_i}$ sobre el cubo I^n. Podemos integrar esta función primero respecto de la i-ésima variable x_i (entre 0 y 1). Es claro que el resultado de la integral indefinida será la función f_i, la cual se deberá evaluar (*en su i-ésima variable*) entre 0 y 1, resultando entonces la diferencia

$$f_i(x_1,\ldots,x_{i-1},1,x_i,\ldots,x_n)-f_i(x_1,\ldots,x_{i-1},0,x_i,\ldots,x_n)$$

Nótese que cada uno de estos sumandos no es más que la composición de la función f con las funciones ∂_i, ∂_i': $I^{n-1} \to I^n$

$$\partial_i(x_1, x_2, \ldots, x_{n-1}) = (x_1, \ldots, x_{i-1}, 0, x_i, \ldots, x_{n-1})$$
$$\partial_i'(x_1, x_2, \ldots, x_{n-1}) = (x_1, \ldots, x_{i-1}, 1, x_i, \ldots, x_{n-1})$$

definidas en la sección anterior. En forma más precisa, se tiene

$$f_i(x_1, \ldots, x_{i-1}, 1, x_i, \ldots, x_n) = (f_i \circ \partial_i')(x_1, x_2, \ldots, x_{n-1})$$
$$f_i(x_1, \ldots, x_{i-1}, 0, x_i, \ldots, x_n) = (f_i \circ \partial_i)(x_1, x_2, \ldots, x_{n-1})$$

Así pues,

$$\begin{aligned}\int_c d\omega &= \sum_{i=1}^{n}(-1)^{i-1}\int_{I^{n-1}}\left(\int_0^1 \frac{\partial f_i}{\partial x_i}dx_i\right)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n \\ &= \sum_{i=1}^{n}(-1)^{i-1}\int_{I^{n-1}}(f_i \circ \partial_i' - f_i \circ \partial_i)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n\end{aligned}$$

Nótese que (viendo a las funciones ∂_i, ∂_i': $I^{n-1} \to I^n$ como $(n-1)$-cubos singulares en $\mathbb{R}^n$)

$$\begin{aligned}\int_{\partial_i'} \omega &= \int_{I^{n-1}} \partial_i'^* \omega = \int_{I^{n-1}} \partial_i'^*\left(\sum_{j=1}^{n} f_j dx_1 \cdots dx_{j-1}dx_{j+1}\cdots dx_n\right) \\ &= \int_{I^{n-1}} \sum_{j=1}^{n} \partial_i'^*(f_j)\partial_i'^*(dx_1)\cdots \partial_i'^*(dx_{j-1})\partial_i'^*(dx_{j+1})\cdots \partial_i'^*(dx_n)\end{aligned}$$

Recuerde que

$$(\partial_i^*)(dx_j) = (\partial_i'^*)(dx_j) = \begin{cases} dx_j & \text{si } j < i \\ 0 & \text{si } j = i \\ dx_{j-1} & \text{si } j > i \end{cases}$$

(ver teorema 10.5.3 en la sección anterior). Obsérvese entonces que si $j < i$ o bien si $j > i$, en la expresión $dx_1 \cdots dx_{j-1}dx_{j+1} \cdots dx_n$ deberá aparecer un dx_i. En tal caso $\partial_i'^*(dx_i) = 0$. Por lo tanto, el único sumando que sobrevive es el que corresponde a $j = i$ (este es el único sumando donde *no aparece* dx_i). Así pues

$$\int_{\partial_i'} \omega = \int_{I^{n-1}} \partial_i'^*(f_i) = \int_{I^{n-1}} f_i \circ \partial_i'$$

Un argumento análogo muestra que

$$\int_{\partial_i} \omega = \int_{I^{n-1}} f_i \circ \partial_i$$

o bien, en términos de integrales $(n-1)$-múltiples

$$\int_{\partial_i'} \omega = \int_{I^{n-1}} f_i \circ \partial_i' dx_1 \cdots dx_{i-1}dx_{i+1} \cdots dx_n$$
$$\int_{\partial_i} \omega = \int_{I^{n-1}} f_i \circ \partial_i dx_1 \cdots dx_{i-1}dx_{i+1} \cdots dx_n$$

En resumen,

$$\begin{aligned}\int_c d\omega &= \sum_{i=1}^{n}(-1)^{i-1}\int_{I^{n-1}}(f_i \circ \partial_i' - f_i \circ \partial_i)dx_1 \cdots dx_{i-1}dx_{i+1}\cdots dx_n \\ &= \sum_{i=1}^{n}(-1)^{i-1}\left(\int_{\partial_i'}\omega - \int_{\partial_i}\omega\right) = \sum_{i=1}^{n}(-1)^{i}\left(\int_{\partial_i}\omega - \int_{\partial_i'}\omega\right)\end{aligned}$$

Ahora bien, recuerde que la n-cadena singular c que estamos considerando es la formada por el n-cubo singular $T\colon I^n \to \mathbb{R}^n$ dado por $T(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n)$. Entonces las funciones $\partial_i, \partial_i'\colon I^{n-1} \to I^n$ se pueden escribir como $T \circ \partial_i$ y $T \circ \partial_i'$, respectivamente. Así pues, se tiene

$$\begin{aligned}\int_c d\omega &= \sum_{i=1}^{n}(-1)^{i}\left(\int_{T\circ\partial_i}\omega - \int_{T\circ\partial_i'}\omega\right) = \sum_{i=1}^{n}(-1)^{i}\int_{T\circ\partial_i - T\circ\partial_i'}\omega \\ &= \int_{\sum_{i=1}^{n}(-1)^i(T\circ\partial_i - T\circ\partial_i')}\omega\end{aligned}$$

expresión que no es más que la integral de la forma ω sobre la frontera de la n-cadena c (es decir, la frontera del n-cubo T, definida justamente como $\partial T = \sum_{i=1}^{n}(-1)^i(T \circ \partial_i - T \circ \partial_i')$). Entonces

$$\int_c d\omega = \int_{\partial c}\omega$$

lo que prueba el teorema en este caso particular (en el que ω es una $(n-1)$- forma y c es la n-cadena singular formada por el único n-cubo singular "identidad").

Consideremos ahora el caso más en general de la $(p-1)$-forma ω definida en $U \subseteq \mathbb{R}^n$ y c la p-cadena singular en U formada por el único p-cubo singular $T\colon I^p \to U$. Se tiene

$$\int_T d\omega = \int_{I^p} T^*(d\omega) = \int_{I^p} d(T^*\omega)$$

(teorema 10.3.3). Obsérvese que $T^*\omega$ es una $(p-1)$-forma definida en (un abierto que contiene a) $I^p \subset \mathbb{R}^p$. Consideremos el p-cubo singular identidad del caso anterior, que denotaremos por Id: $I^p \to \mathbb{R}^p$. Por el resultado previamente probado, se tiene

$$\int_{I^p} d(T^*\omega) = \int_{\mathrm{Id}} d(T^*\omega) = \int_{\partial \mathrm{Id}} T^*\omega$$

donde $\partial\,\mathrm{Id} = \sum_{i=1}^{p}(-1)^i(\mathrm{Id}\circ\partial_i - \mathrm{Id}\circ\partial_i') = \sum_{i=1}^{p}(-1)^i(\partial_i - \partial_i')$. Entonces

$$\begin{aligned}\int_{\partial\,\mathrm{Id}} T^*\omega &= \sum_{i=1}^{p}(-1)^i(\partial_i - \partial_i')^*(T^*\omega) = \sum_{i=1}^{p}(-1)^i(\partial_i^*(T^*\omega) - \partial_i'^*(T^*\omega)) \\ &= \sum_{i=1}^{p}(-1)^i((T\circ\partial_i)^*\omega - (T\circ\partial_i'^*)\omega) = \sum_{i=1}^{p}(-1)^i((T\circ\partial_i)^* - (T\circ\partial_i'^*))\omega\end{aligned}$$

(ver teorema 10.3.2). Esta última expresión es justamente la integral de la $(p-1)$-forma ω sobre la frontera de T (es decir, de c). Entonces hemos probado que para este caso (c un p-cubo arbitrario) se tiene también que

$$\int_c d\omega = \int_{\partial c} \omega$$

Supongamos por último que c es una p-cadena singular arbitraria en $U \subseteq \mathbb{R}^n$, digamos $c = a_1T_1 + a_2T_2 + \cdots + a_kT_k$, en donde $T_i \colon I^p \to U$ son p-cubos singulares en U. Se tiene (con el caso ya demostrado)

$$\int_{\partial c} \omega = \sum_{i=1}^{k} a_i \int_{\partial T_i} \omega = \sum_{i=1}^{k} a_i \int_{T_i} d\omega = \int_c d\omega$$

Así queda plenamente probado este teorema. Q.E.D.

Termina esta sección (y el libro) viendo que dos de los grandes resultados a los que se llegó en los capítulos previos al presente, a saber, el teorema de Green (capítulo 7, sección 9) y el teorema de Stokes (capítulo 9, sección 5), resultan ser casos particulares del teorema probado aquí: si ω una $(p-1)$-forma en $U \subseteq \mathbb{R}^n$, y c una p-cadena singular en U, entonces

$$\int_{\partial c} \omega = \int_c d\omega$$

Advertimos que no vamos a ser tan minuciosos ni detallistas como cuando se estudiaron los resultados mencionados, por ejemplo, respecto a la orientación de las curvas (en el teorema de Green) y a la orientación de superficies (en el teorema de Stokes). Sólo se harán algunas operaciones que nos permitan ver de manera rápida que en efecto las fórmulas de los teoremas de Green y Stokes están contenidas en la fórmula del teorema 10.6.1.

Veamos, a manera de preámbulo, el caso de $p = n = 1$. Se tiene una 0-forma ω en $U \subseteq \mathbb{R}$, digamos $\omega = f \colon U \subseteq \mathbb{R} \to \mathbb{R}$, una función real de una variable real definida en el conjunto abierto U de $\mathbb{R}$, y c una 1-cadena singular en U. Aún más, supongamos que c está formada por el único 1-cubo singular $T \colon [0, 1] \to U$, $T(t) = t$. En este caso la frontera de c es la 0-cadena $\partial c = T_1 - T_2$, en donde $T_1, T_2 \colon \{0\} \to U$ son los 0-cubos singulares $T_1(0) = 1$, $T_2(0) = 0$ (verifique). La fórmula demostrada en el teorema 10.6.1 toma entonces el aspecto

$$\begin{aligned}\int_{[0,1]} f'(x)dx &= \int_c d\omega = \int_{\partial c} \omega \\ &= \int_{T_1 - T_2} f = \int_{T_1} f - \int_{T_2} f = f(T_1(0)) - f(T_2(0)) = f(1) - f(0)\end{aligned}$$

o bien, con notación más familiar

$$\int_0^1 f'(x)dx = f(1) - f(0)$$

que no es más que nuestro viejo conocido teorema fundamental del cálculo del primer curso de cálculo diferencial e integral de funciones reales de una variable real.

Antes de ver los siguientes casos, hagamos una observación en general: sea $T\colon I^n \to \mathbb{R}^n$ es un n-cubo singular, y sea $R = T(I^n)$. Suponga que η es una n-forma en el conjunto abierto $U \subseteq \mathbb{R}^n$, el cual contiene a R. Digamos que $\eta = f dx_1 dx_2 \cdots dx_n$. Entonces

$$\int_T \eta = \int_{I^n} T^*\eta = \int_{I^n} T^*(f dx_1 dx_2 \cdots dx_n) = \int_{I^n} (f \circ T)(\det T'(x)) dx_1 dx_2 \cdots dx_n$$

(ver teorema 10.3.4). Nótese que la última integral (que es una integral n-múltiple de la función $f \circ T$ sobre I^n) no es más que la expresión que se obtiene en la fórmula de cambio de variables de la integral n-múltiple de la función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ sobre la región R. Entonces podemos escribir

$$\int_T \eta = \int_R f dx_1 dx_2 \cdots dx_n$$

Consideremos entonces el caso particular del teorema 10.6.1 cuando $p = 2$. Sea ω una 1-forma definida en el conjunto abierto $U \subseteq \mathbb{R}^2$, digamos $\omega = f_1 dx_1 + f_2 dx_2$. Sea $c = T$ la 2-cadena singular en U constituída por el único 2-cubo singular $T\colon I^2 \to \mathbb{R}^2$. Sea $R = T(I^2)$. La frontera de T (que es un 1-cubo singular $\partial T\colon I^1 \to \mathbb{R}^2$) se puede pensar como una parametrización de la frontera de la región R = imagen de T. Escribamos $\lambda = \partial T$. Entonces, el teorema 10.6.1 nos dice que

$$\int_\lambda \omega = \int_\lambda f_1 dx_1 + f_2 dx_2 = \int_{\partial c} \omega = \int_c d\omega$$

donde

$$\begin{aligned} d\omega &= df_1 dx_1 + df_2 dx_2 \\ &= \left(\frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2\right) dx_1 + \left(\frac{\partial f_2}{\partial x_1} dx_1 + \frac{\partial f_2}{\partial x_2} dx_2\right) dx_2 \\ &= \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 dx_2 \end{aligned}$$

Nótese que $d\omega$ es una 2-forma definida en $\mathbb{R}^2$. La integral de esta 2-forma sobre el 2-cubo singular $T\colon I^2 \to \mathbb{R}^2$ (es decir, sobre la 2-cadena singular c) es

$$\begin{aligned} \int_c d\omega &= \int_{I^2} T^*(d\omega) = \int_{I^2} T^*\left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 dx_2 \\ &= \int_R\int \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 dx_2 \end{aligned}$$

(por la observación hecha previamente a la exposición de este caso). Entonces la fórmula del teorema 10.6.1 se ve en este caso como

$$\int_\lambda f_1 dx_1 + f_2 dx_2 = \int_R\int \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 dx_2$$

que no es más que la fórmula del teorema de Green.

Sea ahora $T\colon I^2 \to \mathbb{R}^3$ la parametrización de una superficie simple (capítulo 8, sección 1), que veremos como una 2-cadena singular en $\mathbb{R}^3$ formada por el único 2-cubo singular T en $\mathbb{R}^3$. Consideremos una 1-forma ω definida en el conjunto abierto $U \subseteq \mathbb{R}^3$, el cual contiene a $S = T(I^2)$,

digamos $\omega = f_1 dx_1 + f_2 dx_2 + f_3 dx_3$. Nuevamente, la frontera de T (que es un 1-cubo singular $\mu = \partial T: I^1 \to \mathbb{R}^3$) puede ser pensada como una parametrización de la curva frontera de la superficie S. Escribamos ∂S para denotar dicha curva. Entonces

$$\int_c \omega = \int_{\partial T} \omega = \int_{\partial S} \mathbf{F} \cdot d\boldsymbol{\mu}$$

Donde $\mathbf{F} = (f_1, f_2, f_3)$ es el campo asociado a la forma ω. Por otra parte

$$\int_c d\omega = \int_{I^2} T^*(d\omega)$$

en que

$$\begin{aligned} d\omega &= df_1 dx_1 + df_2 dx_2 + df_3 dx_3 \\ &= \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right) dx_2 dx_3 - \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\right) dx_1 dx_3 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) dx_1 dx_2 \end{aligned}$$

(ver sección 2), y, si escribimos $T = (T_1, T_2, T_3)$, en donde $T_1, T_2, T_3: I^2 \to \mathbb{R}$, tenemos también que (llamando u y v a las coordenadas de $\mathbb{R}^2$ en donde está I^2 = dominio de T)

$$\begin{aligned} T^*(d\omega) &= T^*\left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right) T^*(dx_2) T^*(dx_3) - T^*\left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\right) T^*(dx_1) T^*(dx_3) \\ &\quad + T^*\left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right) T^*(dx_1) T^*(dx_2) \\ &= \left(\frac{\partial f_3}{\partial x_2} \circ T - \frac{\partial f_2}{\partial x_3} \circ T\right)\left(\frac{\partial T_2}{\partial u} du + \frac{\partial T_2}{\partial v} dv\right)\left(\frac{\partial T_3}{\partial u} du + \frac{\partial T_3}{\partial v} dv\right) \\ &\quad - \left(\frac{\partial f_1}{\partial x_3} \circ T - \frac{\partial f_3}{\partial x_1} \circ T\right)\left(\frac{\partial T_1}{\partial u} du + \frac{\partial T_1}{\partial v} dv\right)\left(\frac{\partial T_3}{\partial u} du + \frac{\partial T_3}{\partial v} dv\right) \\ &\quad + \left(\frac{\partial f_2}{\partial x_1} \circ T - \frac{\partial f_1}{\partial x_2} \circ T\right)\left(\frac{\partial T_1}{\partial u} du + \frac{\partial T_1}{\partial v} dv\right)\left(\frac{\partial T_2}{\partial u} du + \frac{\partial T_2}{\partial v} dv\right) \\ &= \left(\frac{\partial f_3}{\partial x_2} \circ T - \frac{\partial f_2}{\partial x_3} \circ T\right)\left(\frac{\partial T_2}{\partial u}\frac{\partial T_3}{\partial v} - \frac{\partial T_3}{\partial u}\frac{\partial T_2}{\partial v}\right) du dv \\ &\quad + \left(\frac{\partial f_1}{\partial x_3} \circ T - \frac{\partial f_3}{\partial x_1} \circ T\right)\left(\frac{\partial T_1}{\partial v}\frac{\partial T_3}{\partial u} - \frac{\partial T_1}{\partial u}\frac{\partial T_3}{\partial v}\right) du dv \\ &\quad + \left(\frac{\partial f_2}{\partial x_1} \circ T - \frac{\partial f_1}{\partial x_2} \circ T\right)\left(\frac{\partial T_1}{\partial u}\frac{\partial T_2}{\partial v} - \frac{\partial T_1}{\partial v}\frac{\partial T_2}{\partial u}\right) du dv \\ &= (\operatorname{rot} \mathbf{F} \circ T) \cdot \left(\frac{\partial T}{\partial u} \times \frac{\partial T}{\partial v}\right) du dv \end{aligned}$$

Entonces

$$\int_{I^2} T^*(d\omega) = \int_{I^2} (\operatorname{rot} \mathbf{F} \circ T) \cdot \left(\frac{\partial T}{\partial u} \times \frac{\partial T}{\partial v}\right) du dv$$

Esta última integral es la definición de integral de superficie del campo rot **F** sobre la superficie T, que escribimos $\iint_T \operatorname{rot}\mathbf{F}\,dA$. Juntando las piezas, obtenemos la fórmula

$$\int_{\partial S} \mathbf{F} \cdot d\mu = \iint_T \operatorname{rot}\mathbf{F}\,dA$$

que no es más que el teorema de Stokes que estudiamos en la sección 5 del capítulo 9.

Así, el teorema 10.6.3 proporciona un resultado que generaliza a los teoremas más importantes del cálculo en $\mathbb{R}^n$ en los casos $n = 1$ (el llamado teorema fundamental del cálculo), $n = 2$ (el teorema de Green) y $n = 3$ (el teorema de Stokes). Podremos llamarle entonces, con toda propiedad, el TEOREMA FUNDAMENTAL DEL CALCULO EN $\mathbb{R}^n$.

Y consideramos que este resultado nos brinda una bonita manera de finalizar este libro.

Ejercicios (Capítulo 10, Sección 6)

1. Verifique el teorema de Stokes con la 0-forma en $\mathbb{R}^2$, $\omega = x_1x_2$, y la 1-cadena singular $c = T: I^1 \to \mathbb{R}^2$, $T(y) = (y, y^2)$.
2. Verifique el teorema de Stokes con la 1-forma en $\mathbb{R}^3$, $\omega = x_1dx_1 + x_2dx_2 + x_3dx_3$, y la 2-cadena singular $c = T: I^2 \to \mathbb{R}^3$, $T(y_1, y_2) = (y_1y_2, y_1, y_2)$.
3. Verifique el teorema de Stokes con la 3-forma en $\mathbb{R}^4$, $\omega = x_1x_2x_3x_4dx_1dx_3dx_4$, y la 4-cadena singular $c = T_1 + T_2$, donde $T_1, T_2: I^4 \to \mathbb{R}^4$ son los 4-cubos singulares dados por $T_1(y_1, y_2, y_3, y_4) = (y_4, y_3, y_2, y_1)$, $T_2(y_1, y_2, y_3, y_4) = (y_1y_2, y_2y_3, y_3y_4, y_1y_4)$.

Respuestas a los ejercicios

Respuestas a los ejercicios del capítulo 1,

Capítulo 1, Sección 1 (página 12)

2. **a.** $(0, 0, 0)$; **b.** $(-1, -2, 3, -5)$; **c.** v; **d.** $\mathbf{0} \in \mathbb{R}^n$; **e.** $(4, 3, 3)$; **g.** $(6, 2, -2, 1)$; **h.** $\mathbf{0} \in \mathbb{R}^5$; **i.** $(6, 3, 3)$; **j.** $(-5, -5, -5, -5, -5)$; **l.** $(-8, -16, 28)$; **m.** $(-2, -3, 1, 0)$; **n.** $(-19, -24)$; **o.** $(19, 24)$; **p.** $(15, 15, 44)$; **q.** $(16, -75, 13)$; **r.** $(0, 5, 10, 10)$; **s.** $(30, 0, 0, -6)$; **t.** $(-2, -1, 2, 1)$; **u.** $(10, 8, 12)$
7. **a.** l.i.; **b.** l.d.; **c.** l.d.; **d.** l.d.; **e.** l.i.
15. **a.** l.i.; **b.** l.i.; **c.** l.d.; **d.** l.i.; **e.** l.i.
17. **a.** Sí es subespacio; **b.** Sí es subespacio; **c.** No es subespacio; **d.** Sí es subespacio (solamente contiene al vector cero de $\mathbb{R}^3$); **e.** Sí es subespacio (es todo el espacio $\mathbb{R}^3$); **f.** No es subespacio; **g.** Sí es subespacio; **h.** No es subespacio; **i.** No es subespacio (a menos que $a = \mathbf{0}$).

Capítulo 1, Sección 2 (página 22)

6. Son vectores de la forma $t(1, 3)$. Se encuentran sobre la recta $y = 3x$.
8. **a.** $(x, y) = t(1, 2)$; **b.** $(x, y) = (0, 1) + t(1, 1)$; **c.** $(x, y) = (0, 3) + t(1, -2)$; **d.** $(x, y) = (0, -1) + t(1, -1)$
18. Falso.;
19. Falso.;
23. **a.** $(52/29, 130/29)$; **b.** $(4, 0)$; **c.** $(2, 1)$; **d.** $(4/5, 2/5, 0)$; **e.** $(8/7, 8/7, 8/7, 16/7)$.
24. $(21/13, 14/13, 0)$.
25. **a.** $19/2$; **b.** $31/2$; **c.** $35/2$; **d.** $\sqrt{11}/2$.

Capítulo 1, Sección 3 (página 31)

1. **a.** 4; **b.** $\sqrt{10}$; **c.** $\sqrt{6}$; **d.** $\sqrt{29}$; **e.** $\sqrt{8}$; **f.** $\sqrt{15}$.
2. No es inyectiva. Es sobreyectiva.
9. No (atendiendo a la desigualdad triangular).
10. $\|x + y\| = \|x - y\| = \sqrt{58}$.
11. $\|x - y\| = \sqrt{33}$;
12. $\|x + y\| = 20$.
16. **a.** $\arccos(7\sqrt{2}/10)$; **b.** $\arccos(1/\sqrt{14})$; **c.** $\arccos(13\sqrt{58}/174)$.
17. π;
18. **a.** 6; **b.** $\sqrt{37}$; **c.** $\sqrt{13}$;
19. 5.
27. **a.** $4\sqrt{2}$; **b.** $\sqrt{10}$; **c.** $2\sqrt{3}$.
28. En radianes: 1.107, 1.107, 0.927.
29. En radianes: 0.66169, 0.66169, 1.81821.
31. **a.** $(5, 10)$; **b.** $(3, 9/2, 1)$; **c.** $(2, 2, 2, 2)$.

32. $(4, 5), (5/2, 11/2), (9/2, 13/2)$.
33. Los vértices son $A = (1, -3)$, $B = (3, 1)$, $C = (-5, 7)$
34. $5\sqrt{2}/2$;
36. **a.** 2; **b.** $3\sqrt{2}/2$; **c.** 0; **d.** $3\sqrt{17}/34$.

Capítulo 1, Sección 4 (página 43)

1. **a.** $a = -10$; **b.** $a = 5$; **c.** $a = 0$; **d.** a puede ser cualquier real.
2. $(6, -7) = (-2)(3/5, 4/5) + (-9)(-4/5, 3/5)$.
3. $(3, 1, 0) = (4/3)\mathbf{v}_1 + (7/3)\mathbf{v}_2 + (5/3)\mathbf{v}_3$.
4. $(2, 4, 1, 3) = 5\mathbf{v}_1 + \mathbf{v}_2 - 2\mathbf{v}_3$.
5. $P = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}$, Matriz de cambio de base de β' a $\beta = P^t$.
6. $P = \begin{bmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{bmatrix}$, Matriz de cambio de base de β' a $\beta = P^t$
7. $P = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}$, Matriz de cambio de base de β' a $\beta = P^t$
8. $\beta = \{(2\sqrt{5}/5, \sqrt{5}/5), (-\sqrt{5}/5, 2\sqrt{5}/5)\}$
9. $\beta = \{(\sqrt{2}/2, \sqrt{2}/2), (-\sqrt{2}/2, \sqrt{2}/2)\}$
10. $\beta = \{(2\sqrt{5}/5, \sqrt{5}/5), (\sqrt{5}/5, -2\sqrt{5}/5)\}$
11. $\beta = \{(3\sqrt{10}/10, \sqrt{10}/10), (-\sqrt{10}/10, 3\sqrt{10}/10)\}$
12. $\beta = \{(\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3), (\sqrt{6}/6, \sqrt{6}/6, -\sqrt{6}/3), (\sqrt{2}/2, -\sqrt{2}/2, 0)\}$
13. $\beta = \{(\sqrt{6}/6, \sqrt{6}/3, \sqrt{6}/6), (11\sqrt{210}/210, -4\sqrt{210}/105, \sqrt{210}/42),$
$(-3\sqrt{35}/35, -\sqrt{35}/35, \sqrt{35}/7)\}$
14. $\beta = \{(0, \sqrt{2}/2, \sqrt{2}/2), (\sqrt{6}/3, \sqrt{6}/6, -\sqrt{6}/6), (\sqrt{3}/3, -\sqrt{3}/3, \sqrt{3}/3)\}$
15. $\beta = \{(3\sqrt{11}/11, \sqrt{11}/11, \sqrt{11}/11),$
$(7\sqrt{330}/330, -\sqrt{330}/66, -8\sqrt{330}/165), (\sqrt{30}/30, -\sqrt{30}/6, \sqrt{30}/15)\}$
16. $\beta = \{(1/2, 1/2, 1/2, 1/2), (\sqrt{3}/6, \sqrt{3}/6, \sqrt{3}/6, -\sqrt{3}/2), (\sqrt{6}/6, \sqrt{6}/6, -\sqrt{6}/3, 0),$
$(\sqrt{2}/2, -\sqrt{2}/2, 0, 0)\}$
17. $\beta = \{(\sqrt{3}/3, 0, \sqrt{3}/3, -\sqrt{3}/3), (\sqrt{51}/51, \sqrt{51}/17, 4\sqrt{51}/51, 5\sqrt{51}/51),$
$(-5\sqrt{51}/153, 19\sqrt{51}/153, -\sqrt{51}/51, -8\sqrt{51}/153), (-4\sqrt{3}/9, -\sqrt{3}/9, \sqrt{3}/3, -\sqrt{3}/9)\}$
18. $P = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 9 & 0 \end{bmatrix}$, $P^{-1} = \dfrac{1}{10}\begin{bmatrix} 0 & 9 & 1 \\ 0 & -1 & 1 \\ 1 & -17 & -3 \end{bmatrix}$
19. $P = \begin{bmatrix} 1 & 1 & 3 \\ 2 & -1 & 1 \\ 3 & -2 & 1 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} -1 & 7 & -4 \\ -1 & 8 & -5 \\ 1 & -5 & 3 \end{bmatrix}$
20. $P = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 3 & 2 \\ 5 & 0 & -1 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} -3/5 & 2/5 & 1/5 \\ 2 & -1 & 0 \\ -3 & 2 & 0 \end{bmatrix}$

Capítulo 1, Sección 5 (página 56)

1. $\mathbf{u} \times \mathbf{v} = (-2, -2, 2) = -\mathbf{v} \times \mathbf{u}$
2. $\mathbf{u} \times \mathbf{v} = (24, 0, -16) = -\mathbf{v} \times \mathbf{u}$
3. $\mathbf{u} \times \mathbf{v} = (0, 0, 0) = -\mathbf{v} \times \mathbf{u}$

4. $\mathbf{u} \times \mathbf{v} = (0, -1, 1) = -\mathbf{v} \times \mathbf{u}$
5. $\mathbf{u} \times \mathbf{v} = (0, 0, 0) = -\mathbf{v} \times \mathbf{u}$
6. $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (28, -80, -2)$, $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (-21, -91, 16)$
9. **a.** $(0, -11, 11)$, **b.** $(0, -11, 11)$, **c.** $(0, 0, 0)$, **d.** $-2(0, -11, 11)$, **e.** $(0, 121, -121)$
14. 1/2,
15. 15,
16. $4\sqrt{26}$,
17. $4\sqrt{26}$,
18. 56
29. No son coplanares.
30. Son coplanares. Se encuentran en el plano $-x + 6y - 4z = 0$
31. Son coplanares. Se encuentran en el plano $-x + 4y - z = 0$
32. No son coplanares.
33. $2x + 2y + z = 3$,
34. $x + y + 2z = 4$,
35. $x + y + z = 1$
37. $12\sqrt{2}$,
38. $\sqrt{269}$,
39. 5,
40. 15,
41. 103/2
42. 205/2
43. **a.** $(\sqrt{5}, \arctan(1/2), 1)$, **b.** $(\sqrt{10}, -\arctan 3, 5)$ **c.** $(1, 0, 0)$, **d.** $(\sqrt{13}, \arctan(3/2), -1)$
44. **a.** $(2, 0, 1)$, **b.** $(-1, 0, 3)$, **c.** $(3/2, -3\sqrt{3}/2, -2)$, **d.** $(8\sqrt{2}, 8\sqrt{2}, 0)$
45. $z = r$,
46. $r^2 + z^2 = 1$
47. **a.** $z = r^2$, **b.** $z = r^2(2\cos^2\theta + 3\operatorname{sen}^2\theta)$
48. **a.** $(1, 0, \pi/2)$, **b.** $(\sqrt{11}, \arctan(1/3), \arccos(-1/\sqrt{11}))$ **c.** $(\sqrt{2}, \pi/2, \pi/4)$, **d.** $(\sqrt{38}, \arctan(3/2), \arccos(-5/\sqrt{38}))$
49. **a.** $(0, 0, 1)$, **b.** $(0, 2, 0)$, **c.** $(\sqrt{2}/4, \sqrt{6}/4, -\sqrt{2}/2)$, **d.** $(\sqrt{30}/2, -\sqrt{30}/2, 1)$
50. $r \operatorname{sen}\phi = 3$,
51. $r = 2\operatorname{sen}\phi\cos\theta$
52. **a.** $r = 2a\operatorname{sen}\phi\cos\theta$, **b.** $r = 2a\operatorname{sen}\phi\operatorname{sen}\theta$, **c.** $r = 2a\cos\theta$
53. $\phi = \arctan a$ o $\phi = \pi - \arctan a$

Capítulo 1, Sección 6 (página 66)

1. $x + y + z = 0$,
2. $x - 2 = 0$,
3. $2y + 3z = 23$
4. $3x + 2y + 6z = 4$,
5. $-2x - 7y + 4z = -14$
6. $x_0x + y_0y + z_0z = x_0^2 + y_0^2 + z_0^2$
7. **a.** $-2x - 3y + 2z = 7$; **b.** $2x + 3y - 2z = 10$
8. $17x - 22y - 6z = 57$,
9. $x + 34y - 9z = 35$
10. **p** y **q** pertenecen al plano.
11. **p** pertenece al plano, **q** no pertenece al plano

12. **p** y **q** pertenecen al plano.
13. Ninguno de los dos puntos pertenecen al plano.
14. **p** no pertenece al plano, **q** pertenece al plano.
15. Pasa por puntos del tipo $(t, s, 3 - 3t)$, $t, s \in \mathbb{R}$. Vector normal $(3, 0, 1)$
16. Pasa por puntos del tipo $(t, 0, s)$, $t, s \in \mathbb{R}$. Vector normal $(0, 1, 0)$
17. Pasa por puntos del tipo $(t, s, t - s - 5)$, $t, s \in \mathbb{R}$. Vector normal $(1, -1, -1)$
18. Pasa por puntos del tipo $\left(t, s, \dfrac{23 - 3t + 2s}{23}\right)$, $t, s \in \mathbb{R}$. Vector normal $(3, -2, 7)$
19. Ninguno de los dos.
20. Son paralelos.
21. Son paralelos.
22. Son perpendiculares.,
23. Son paralelos.
24. $3x - 2y + z = 7$
25. $4x - y + z = 0$
26. $2x - 13y + 7z = 0$
27. $2y - x = 0$
28. $8x - 13y + 3z = 18$
29. $6x + 4y - 9z = -59$
30. $x/a + y/b + z/c = 1$
35. $d = 5$
36. $d = 13/\sqrt{16}$
37. $d = 14/\sqrt{17}$
38. $d = 3/\sqrt{24}$
39. $d = \dfrac{|D_1 - D_2|}{\sqrt{A_2^2 + B_2^2 + C_2^2}}$
40. $V = \dfrac{8}{14\sqrt{14}}$
42. $x + y - 2z = 2 \pm 2\sqrt{6}$, $2x + z = 5 \pm 2\sqrt{5}$, $x - 5y - 2z = -10 \pm 2\sqrt{30}$
43. $x - 4y + 7z = 0$, $x - 4y + 7z = 66$, $-3x + y + z = 0$, $-3x + y + z = 11$, $x + 2y + z = 0$, $x + 2y + z = 6$.
44. $\mathbf{p} = (19/6, 0, 0)$
45. $\mathbf{p}_1 = (0, -6/13, 0)$, $\mathbf{p}_2 = (0, 6/7, 0)$
46. $\mathbf{p} = (3, 2, 1)$
47. $\mathbf{p}_1 = (4.9705, 0.99075, -0.46886)$, $\mathbf{p}_2 = (4.2402, 4.64224, 0.99173)$, $\mathbf{p}_3 = (1.39279, 0.99075, -2.25772)$, $\mathbf{p}_4 = (0.6625, 4.64224, -0.79712)$, $\mathbf{p}_5 = (-0.9705, 3.00925, 2.46886)$, $\mathbf{p}_6 = (2.6072, 3.00925, 4.25772)$, $\mathbf{p}_7 = (-0.2402, -0.64224, 1.00827)$, $\mathbf{p}_8 = (3.3375, -0.64224, 2.79712)$
48. $\mathbf{p}_1 = (0, 0, 0)$, $\mathbf{p}_2 = (1, 2, 1)$, $\mathbf{p}_3 = (-3, 1, 1)$, $\mathbf{p}_4 = (1, -4, 7)$ $\mathbf{p}_5 = (-2, 3, 2)$, $\mathbf{p}_6 = (-1, -1, 9)$, $\mathbf{p}_7 = (-2, -3, 8)$, $\mathbf{p}_8 = (2, -2, 8)$
49. $x = y = z$
50. $x = 0$, $y = 1 + t$, $z = 0$, $t \in \mathbb{R}$
51. $x = 2 + 3t$, $y = -4 + t$, $z = -7 + 2t$, $t \in \mathbb{R}$
52. $x = 2$, $y = 1 - 4t$, $z = 1 - 6t$, $t \in \mathbb{R}$
53. $x = 3 + 3t$, $y = 4 - 2t$, $z = 7 + t$, $t \in \mathbb{R}$
54. $x = x_0t$, $y = y_0t$, $z = z_0t$, $t \in \mathbb{R}$

55. $x = 3 + 4t,\ y = 9 + 7t,\ z = 7 + 2t,\ t \in \mathbb{R}$
56. $x = 2 + 4t,\ y = 1 - 2t,\ z = 6 + 4t,\ t \in \mathbb{R}$
57. $x = 2t,\ y = 6t,\ z = 5t,\ t \in \mathbb{R}$
58. $L_1: x = 2 - 2.9705t,\ y = 2 + 1.00925t,\ z = 1 + 1.46886t,\ t \in \mathbb{R}$
$L_2: x = 2 - 2.2402t,\ y = 2 - 2.64224t,\ z = 1 + 0.0082t,\ t \in \mathbb{R}$
$L_3: x = 2 + 0.60721t,\ y = 2 + 1.00925t,\ z = 1 + 3.25772t,\ t \in \mathbb{R}$
$L_4: x = 2 + 1.3375t,\ y = 2 - 2.64224t,\ z = 1 + 1.7971t,\ t \in \mathbb{R}$

59. $L_1: x = -t,\ y = -t,\ z = 9t,\ t \in \mathbb{R}$
$L_2: x = -1/2 - 3t/2,\ y = -1/2 - 5t/2,\ z = 9/2 + 7t/2,\ t \in \mathbb{R}$
$L_3: x = -1/2 + 5t/2,\ y = -1/2 - 3t/2,\ z = 9/2 + 7t/2,\ t \in \mathbb{R}$
$L_4: x = -1/2 - 3t/2,\ y = -1/2 + 7t/2,\ z = 9/2 - 5t/2,\ t \in \mathbb{R}$
Se intersectan en $(-1/2, -1/2, 9/2)$

60. **p** pertenece a la recta, **q** no pertenece a la recta.
61. **p** no pertenece a la recta, **q** pertenece a la recta.
62. **p** no pertenece a la recta, **q** pertenece a la recta.
63. $x = 2 + 3t,\ y = 1 + 4t,\ z = 4 - t,\ t \in \mathbb{R}$
64. $x = 2t,\ y = t,\ z = 0,\ t \in \mathbb{R}$
65. Mediana por $A: x = 2 - 4t,\ y = 1 + 8t,\ z = 3 + t,\ t \in \mathbb{R}$
Mediana por $B: x = -2 + 8t,\ y = 7 - 10t,\ z = 5 - 5t,\ t \in \mathbb{R}$
Mediana por $C: x = 2 - 2t,\ y = 3 + t,\ z = 2 + t,\ t \in \mathbb{R}$
Se cortan en $(2/3, 11/3, 10/3)$

66. Intersección con el plano xy: $(19/5, 6/5, 0)$
Intersección con el plano xz: $(5, 0, -6)$
Intersección con el plano yz: $(0, 5, 19)$

67. $\mathbf{p} = (1, -2, -1)$
70. $x = -4/7 + 2t,\ y = 12/7 + t,\ z = 7t,\ t \in \mathbb{R}$
71. $x = 2/5 - t,\ y = -6/5 + 23t,\ z = 5t,\ t \in \mathbb{R}$
72. $x = 5/2 - t,\ y = -1/2,\ z = t,\ t \in \mathbb{R}$
73. $x = y = 0,\ z = t,\ t \in \mathbb{R}$
74. $x - 2y + z = 0$
75. Son rectas del tipo $\dfrac{x-2}{a} = \dfrac{y-1}{b} = \dfrac{z+1}{c}$, donde $a - b + c = 0$
76. Son rectas del tipo $\dfrac{x-1}{a} = \dfrac{y-3}{b} = \dfrac{z-2}{c}$, donde $a + b - 2c = 0$
77. $d = \sqrt{54}$. No es la distancia entre los planos.
78. $d = \sqrt{56}/7$. Sí es la distancia entre los planos.
87. $d = 0$
88. $d = \sqrt{21910}/35$
89. $d = \sqrt{13594}/14$
90. $d = 1/\sqrt{5}$
91. $d = 13\sqrt{238}/119$
92. $d = 7\sqrt{114}/114$

93. **b.** $x + y - z - u = 0$; **c.** $y + u = 1$; **d.** $x + 3y - 6z + 5u + 2 = 0$; **f.** $(1, 2, 3, -1)$; **g.** $x - y - z + u = 0$; **h.** $-6x + 11y + 6z + 2u - 3v = 0$; **i.** $x_1/a_1 + x_2/a_2 + \cdots + x_n/a_n = 1$.

94. **a.** $x = y = z = u$; **b.** $x = 2 - t$, $y = 1 + t$, $z = 1 + 3t$, $u = 4 - 5t$, $t \in \mathbb{R}$; **c.** $x = 1 + 3t$, $y = -1 - 2t$, $z = 3 + 4t$, $u = -4 - 3t$, $t \in \mathbb{R}$; **d.** $x_1 = x_2 = \cdots = x_n$; **e.** $x = 2 + t$, $y = z = 1$, $u = 3 + 3t$, $t \in \mathbb{R}$; **f.** $x_1 = 1 + 3t$, $x_2 = 4t$, $x_3 = -3t$, $x_4 = t$, $x_5 = 1 + 7t$, $t \in \mathbb{R}$; **g.** $(23, 21, 0, 3)$; **i.** $(2, -1, 3, 1)$; **j.** Mediana por A: $x = 1 - 5t/2$, $y = 1 + 2t$, $z = 2 - t/2$, $u = 1 + t$; Mediana por B: $x = 3 - 3t/2$, $y = 1 + 2t$, $z = 3$, $u = -2 - t/2$; Mediana por C: $x = 4 - t$, $y = -3$, $z = 2 - t/2$, $u = 2 + 3t/2$. Punto común de las tres medianas: $(6, -3, 3, -1)$.

Capítulo 1, Sección 7 (página 78)

1. $\operatorname{Ker} T = \{(x, y) \mid y = -x\}$, $\operatorname{Im} T = \mathbb{R}$.

2. $\operatorname{Ker} T = \{(0, 0)\}$, $\operatorname{Im} T = \mathbb{R}^2$.

3. $\operatorname{Ker} T = \{(x, y, z) \mid x = t, y = -t, z = 3t, t \in \mathbb{R}\}$, $\operatorname{Im} T = \mathbb{R}^2$.

4. $\operatorname{Ker} T = \{(x, y, z) \mid x = y = z\}$, $\operatorname{Im} T = \{(x, y, z) \mid x + y + z = 0\}$.

5. $\operatorname{Ker} T = \{(x, y, z, u) \mid x = -t, y = t, z = -s, u = s, t, s \in \mathbb{R}\}$, $\operatorname{Im} T = \mathbb{R}^2$.

8. $T(x, y) = (2x - 3y, 5x + 4y)$.

9. $T(x, y) = (x + 2y, -5x + 7y)$.

10. El plano $y = x$.

18. $[T] = [a]$, $T^{-1}(x) = a^{-1}x$.

19. $[T] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $T^{-1}(x, y) = (x/2 + y/2, x/2 - y/2)$

20. $[T] = \begin{bmatrix} 10 & 6 \\ 1 & -3 \end{bmatrix}$, $T^{-1}(x, y) = (x/12 + y/6, x/36 - 5x/18)$

21. $[T] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$, $T^{-1}(x, y, z) = (x, -x + y, -y + z)$

22. $[T] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $T^{-1}(x, y, z) = (y, z, x)$

35. **d.** $(-5, 5, 5, -5)$; **e.** $x - y - z + u = 0$; **f.** $(-6, 11, 6, 2, -3)$; **g.** $-6x + 11y + 6z + 2u - 3v = 0$

Capítulo 1, Sección 8 (página 88)

1. polinomio característico $\lambda^2 - 5\lambda + 6$, valores propios 2, 3; vectores propios $t(1, 5)$, $t(2, 5)$

2. polinomio característico $\lambda^2 - 7\lambda + 6$, valores propios 1, 6; vectores propios $t(2, -1)$, $t(3, 1)$

3. polinomio característico $\lambda^2 - 8\lambda + 15$, valores propios 3, 5; vectores propios $t(1, 1)$, $t(2, 1)$

4. polinomio característico $\lambda^2 - 3\lambda - 4$, valores propios -1, 4; vectores propios $t(2, -1)$, $t(1, -3)$

5. polinomio característico $\lambda^2 - 12\lambda + 27$, valores propios 3, 9; vectores propios $t(1, 5)$, $t(7, 5)$

6. polinomio característico $\lambda^2 + 2\lambda - 3$, valores propios 1, -3; vectores propios $t(1, 1)$, $t(1, -1)$

7. polinomio característico $\lambda^2 - 12\lambda + 35$, valores propios 5, 7; vectores propios $t(2, 1)$, $t(4, 1)$

8. polinomio característico $\lambda^2 + 7\lambda + 12$, valores propios -3, -4; vectores propios $t(1, -2)$, $t(1, -1)$

9. polinomio característico $\lambda^2 - 4\lambda + 3$, valores propios 1, 3; vectores propios $t(0, 1)$, $t(1, 1)$

10. polinomio característico $\lambda^2 - 13\lambda + 42$, valores propios 6, 7; vectores propios $t(1, 0)$, $t(0, 1)$

11. polinomio característico $\lambda^2 - 5\lambda + 4$, valores propios 1, 4; vectores propios $t(1, 0)$, $t(1, 3)$

12. polinomio característico $\lambda^2 + 5\lambda + 4$, valores propios -1, -4; vectores propios $t(1, 1)$, $t(2, -1)$

13. polinomio característico $-\lambda^3 + 6\lambda^2 - 11\lambda + 6$, valores propios 1, 2, 3; vectores propios $t(1, 1, -2)$, $t(2, 1, -2)$, $t(1, 1, 0)$

14. polinomio característico $-\lambda^3 + 5\lambda^2 - 2\lambda - 8$, valores propios -1, 2, 4; vectores propios $t(5, 8, 6)$, $t(20, 11, 15)$, $t(20, 7, 9)$

15. polinomio característico $-\lambda^3 + \lambda^2 + 10\lambda + 8$, valores propios -1, -2, 4; vectores propios $t(1, -1, -4)$, $t(1, -2, -5)$, $t(4, 1, 4)$

16. polinomio característico $-\lambda^3 + 8\lambda^2 + 13\lambda - 140$, valores propios -4, 5, 7; vectores propios $t(2, -1, 3)$, $t(14, 11, -6)$, $t(7, 2, -6)$

17. polinomio característico $-\lambda^3 + 12\lambda^2 - 21\lambda + 18$, valores propios 1, 10; vectores propios $t(27, -4, -24)$, $t(27, 50, 48)$

18. polinomio característico $-\lambda^3 + 6\lambda^2 - 12\lambda + 8$, valores propios 2; vectores propios $t(14, -15, -27)$

19. polinomio característico $-\lambda^3 + 2\lambda^2 + 3\lambda$, valores propios 0, -1, 3; vectores propios $t(5, 6, 8)$, $t(1, 2, 2)$, $t(5, 2, 2)$

20. polinomio característico $-\lambda^3 - 10\lambda^2 - 31\lambda - 30$, valores propios -2, -3, -5; vectores propios $t(1, -7, -16)$, $t(1, -6, -8)$, $t(1, -10, 8)$

21. polinomio característico $-\lambda^3 + 9\lambda^2 - 26\lambda + 24$, valores propios 2, 3, 4; vectores propios $t(28, -3, -18)$, $t(14, -3, -10)$, $t(28, -9, -24)$

22. polinomio característico $-\lambda^3 + 5\lambda^2 - 2\lambda - 8$, valores propios -1, 2, 4; vectores propios $t(1, 3, -1)$, $t(2, -3, 1)$, $t(2, -9, -7)$

23. polinomio característico $-\lambda^3 + 2\lambda^2 + 3\lambda$, valores propios 0, -1, 3; vectores propios $t(4, 2, -5)$, $t(1, 1, -1)$, $t(1, -1, 1)$

24. polinomio característico $(\lambda - 1)^3$, valores propios 1; vectores propios $t(5, 8, 20)$

25. polinomio característico $-\lambda^3 - 3\lambda^2 + \lambda + 3$, valores propios 1, -1, -3; vectores propios $t(7, 0, -1)$, $t(7, 4, -3)$, $t(21, 24, -23)$

26. polinomio característico $-\lambda^3 + 6\lambda^2 - 11\lambda + 6$, valores propios 1, 2, 3; vectores propios $t(7, 18, 42)$, $t(35, 60, 126)$, $t(7, 6, 14)$

27.

valor propio	vectores propios $t \in \mathbb{R}$, $t \neq 0$
1	$t(3, 1, 0, 0)$
-1	$t(1, 1, 0, 0)$
2	$t(10, -5, -9, -3)$
-1	$t(2, 3, -1, 1)$

37. $P = \begin{bmatrix} 1 & 2 \\ 5 & 5 \end{bmatrix}$;

38. $P = \begin{bmatrix} 2 & 1 \\ -1 & -3 \end{bmatrix}$

39. $P = \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$;

40. $P = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}$

41. $P = \begin{bmatrix} 1 & 1 & 4 \\ -1 & -2 & 1 \\ -4 & -5 & 4 \end{bmatrix}$,

42. $P = \begin{bmatrix} 5 & 1 & 5 \\ 6 & 2 & 2 \\ 8 & 2 & 2 \end{bmatrix}$

43. a. $A^n = \frac{1}{5}\begin{bmatrix} (3)(6^n)+2 & 6^{n+1}-6 \\ 6^n-1 & (2)(6^n)+3 \end{bmatrix}$

b. $A^n = \begin{bmatrix} (2)(5^n)-3^n & 2(3^n)-2(5^n) \\ 5^n-3^n & 2(3^n)-5^n \end{bmatrix}$

c. $A^n = \begin{bmatrix} (2)(-4^n)-(-3^n) & (-4)^n-(-3)^n \\ 2(-3)^n-2(-4)^n & 2(-3)^n-(-4)^n \end{bmatrix}$

Capítulo 1, Sección 9 (página 98)

1. $q(x, y) = 3x^2 + 5y^2 + 4xy$
2. $q(x, y) = x^2 + 2y^2$
3. $q(x, y) = 4xy$
4. $q(x, y, z) = 2x^2 + 5y^2 + 8z^2 + 2xy + 8xz + 6yz$
5. $q(x, y, z) = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz$
6. $q(x, y, z) = 3x^2 + 5y^2 + 9z^2 + 6xz + 6yz$
7. $q(x, y, z) = 7x^2 - 5y^2 + 2z^2 - 2xy - 6xz$
8. $q(x_1, x_2, x_3, x_4) = 3x_1^2 + 5x_2^2 + 7x_3^2 + 8x_4^2 + 4x_1x_2 - 2x_1x_4 + 8x_2x_3 - 4x_2x_4 - 14x_3x_4$
9. $A = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$;
10. $A = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}$
11. $A = \begin{bmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix}$;
12. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{bmatrix}$
13. $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -5 \\ 2 & -5 & 0 \end{bmatrix}$;
14. $A = \begin{bmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 \end{bmatrix}$
15. $A = \begin{bmatrix} 1 & 0 & -1/2 & 0 \\ 0 & 2 & 0 & -1/2 \\ -1/2 & 0 & 3 & 0 \\ 0 & -1/2 & 0 & 4 \end{bmatrix}$
16. $\Delta_1 = 2, \Delta_2 = 5$; la matriz es definida positiva
17. $\Delta_1 = 3, \Delta_2 = 6$; la matriz es definida positiva
18. $\Delta_1 = -1, \Delta_2 = -10$; la matriz es indefinida
19. $\Delta_1 = -3, \Delta_2 = 7$; la matriz es definida negativa
20. $\Delta_1 = 2, \Delta_2 = 8, \Delta_3 = 66$; definida positiva
21. $\Delta_1 = -2, \Delta_2 = 12, \Delta_3 = -216$; definida negativa
22. $\Delta_1 = 2, \Delta_2 = 16, \Delta_3 = -1$; indefinida
23. $\Delta_1 = -2, \Delta_2 = 4, \Delta_3 = -8$; definida negativa
24. $\Delta_1 = 2, \Delta_2 = 1, \Delta_3 = 8$; definida positiva
25. $\Delta_1 = 3, \Delta_2 = 6, \Delta_3 = 12$; definida positiva

26. $\Delta_1 = -5, \Delta_2 = 3, \Delta_3 = 256/5$; indefinida
27. $\Delta_1 = -1, \Delta_2 = -5, \Delta_3 = -10$; indefinida
28. $\Delta_1 = -3, \Delta_2 = 8, \Delta_3 = 4/3$; indefinida
29. $\Delta_1 = -2, \Delta_2 = 12, \Delta_3 = -40$; definida negativa
30. $\Delta_1 = -1, \Delta_2 = 4, \Delta_3 = -32$; definida negativa
31. $\Delta_1 = 4, \Delta_2 = 1, \Delta_3 = 36$; definida positiva
32. $\Delta_1 = 3, \Delta_2 = -9, \Delta_3 = -18$; indefinida
33. $\Delta_1 = -1, \Delta_2 = 9, \Delta_3 = 72$; indefinida
34. $\Delta_1 = -3, \Delta_2 = -10, \Delta_3 = -6$; indefinida
35. $\Delta_1 = -3, \Delta_2 = 36, \Delta_3 = -33$; definida negativa
36. $\Delta_1 = -3, \Delta_2 = 320, \Delta_3 = -2533$; definida negativa
37. $\Delta_1 = 1, \Delta_2 = 1, \Delta_3 = 4$; definida positiva
38. $\Delta_1 = 4, \Delta_2 = -5, \Delta_3 = 1$; indefinida
39. $\Delta_1 = -1, \Delta_2 = 27, \Delta_3 = -53$; definida negativa
40. $\Delta_1 = 8, \Delta_2 = 128, \Delta_3 = -578/9$; indefinida
41. $\Delta_1 = -3, \Delta_2 = 3, \Delta_3 = 174$; indefinida
42. $\Delta_1 = -4, \Delta_2 = 48, \Delta_3 = -12$; definida negativa
43. $\Delta_1 = 1, \Delta_2 = 4, \Delta_3 = 1$; definida positiva
44. $\Delta_1 = 1, \Delta_2 = 1, \Delta_3 = 5$; definida positiva
45. $\Delta_1 = -4, \Delta_2 = 12, \Delta_3 = -16$; definida negativa
46. $\Delta_1 = -2, \Delta_2 = -2, \Delta_3 = -2$; indefinida
47. $\Delta_1 = -5, \Delta_2 = 16, \Delta_3 = -27$; definida negativa
48. $\Delta_1 = 4, \Delta_2 = 24, \Delta_3 = 168$; definida positiva
49. $\Delta_1 = -2, \Delta_2 = 5, \Delta_3 = -12$; definida negativa
50. $\Delta_1 = -2, \Delta_2 = 16, \Delta_3 = -14$; definida negativa
51. valores propios 5, 10, definida positiva; $5(4x/5 - 3y/5)^2 + 10(3x/5 + 4y/5)^2$
52. valores propios 5, 25, definida positiva; $5(4x/5 - 3y/5)^2 + 25(3x/5 + 4y/5)^2$
53. valores propios 13, 169, definida positiva; $13(-12x/13 + 5y/13)^2 + 169(5x/13 + 12y/13)^2$
54. valores propios 13, 26, definida positiva; $13(-12x/13 + 5y/13)^2 + 26(5x/13 + 12y/13)^2$
55. valores propios 25, 125, definida positiva; $25(4x/5 - 3y/5)^2 + 125(3x/5 + 4y/5)^2$
56. valores propios -5, -75, definida negativa; $-5(4x/5 - 3y/5)^2 - 75(3x/5 + 4y/5)^2$
57. valores propios -13, -39, definida negativa; $-13(-12x/13+5y/13)^2 - 39(5x/13+12y/13)^2$
58. valores propios 1, 4, definida positiva; $(-12x/13 + 5y/13)^2 + 4(5x/13 + 12y/13)^2$
59. valores propios 13, -13, indefinida; $13(-12x/13 + 5y/13)^2 - 13(5x/13 + 12y/13)^2$
60. valores propios -50, -100, definida negativa; $-50(4x/5 - 3y/5)^2 - 100(3x/5 + 4y/5)^2$
61. valores propios 3, 6, 9, definida positiva
$3(2x/3 - 2y/3 + z/3)^2 + 6(2x/3 + y/3 - 2z/3)^2 + 9(x/3 + 2y/3 + 2z/3)^2$
62. valores propios -9, -27, -18, definida negativa;
$-9(2x/3 - 2y/3 + z/3)^2 - 27(2x/3 + y/3 - 2z/3)^2 - 18(x/3 + 2y/3 + 2z/3)^2$
63. valores propios 7, 21, 49, definida positiva;
$7(2x/7 - 3y/7 - 6z/7)^2 + 21(-6x/7 + 2y/7 - 3z/7)^2 + 49(3x/7 + 6y/7 - 2z/7)^2$
64. valores propios 7, 14, 21, definida positiva;
$7(2x/7 - 3y/7 - 6z/7)^2 + 14(-6x/7 + 2y/7 - 3z/7)^2 + 21(3x/7 + 6y/7 - 2z/7)^2$
65. valores propios 5, 10, 15, definida positiva; $5z^2 + 10(4x/5 - 3y/5)^2 + 15(3x/5 + 4y/5)^2$
66. valores propios 9, 18, 81, definida positiva;
$9(-x/9 - 4y/9 - 8z/9)^2 + 18(4x/9 + 7y/9 - 4z/9)^2 + 81(-8x/9 + 4y/9 - z/9)^2$
67. valores propios 5, 25, 50, definida positiva; $5z^2 + 25(4x/5 - 3y/5)^2 + 50(3x/5 + 4y/5)^2$

68. valores propios 6, 9, 12, definida positiva;
$6(2x/3 - 2y/3 + z/3)^2 + 9(2x/3 + y/3 - 2z/3)^2 + 12(x/3 + 2y/3 + 2z/3)^2$
69. valores propios 1, 2, 3, definida positiva;
$(2x/7 - 3y/7 - 6z/7)^2 + 2(-6x/7 + 2y/7 - 3z/7)^2 + 3(3x/7 + 6y/7 - 2z/7)^2$
70. valores propios -1, -7, -14, definida negativa;
$-(2x/7 - 3y/7 - 6z/7)^2 - 7(-6x/7 + 2y/7 - 3z/7)^2 - 14(3x/7 + 6y/7 - 2z/7)^2$
71. valores propios -6, -12, -18 definida negativa;
$-6(2x/3 - 2y/3 + z/3)^2 - 12(2x/3 + y/3 - 2z/3)^2 + 18(x/7 + 2y/7 + 2z/7)^2$
72. valores propios 9, 0, 18, semidefinida positiva;
$9(2x/3 - 2y/3 + z/3)^2 + 18(x/3 + 2y/3 + 2z/3)^2$
73. valores propios $-5, -15, -20$ definida negativa; $-5z^2 - 15(4x/5 - 3y/5)^2 - 20(3x/5 + 4y/5)^2$
74. valores propios 13, 26, 39, definida positiva; $13(-3x/13 - 4y/13 - 12z/13)^2 + 26(12x/13 + 3y/13 - 4z/13)^2 + 39(4x/13 - 12y/13 + 3z/13)^2$
75. valores propios 5, 7, 9, definida positiva; $5(-3x/13 - 4y/13 - 12z/13)^2 + 7(12x/13 + 3y/13 - 4z/13)^2 + 9(4x/13 - 12y/13 - 3z/13)^2$
76. valores propios 15, 9, 30, definida positiva;
$15(-2x/15 - y/3 - 14z/15)^2 + 9(2x/3 + 2y/3 - z/3)^2 + 30(-11x/15 + 10y/15 - 2z/15)^2$
77. valores propios 27, 81, 162, definida positiva;
$27(-x/9 - 4y/9 - 8z/9)^2 + 81(4x/9 + 7y/9 - 4z/9)^2 + 162(-8x/9 + 4y/9 - z/9)^2$
78. valores propios 1, 2, 9, definida positiva;
$(-x/9 - 4y/9 - 8z/9)^2 + 2(4x/9 + 7y/9 - 4z/9)^2 + 9(-8x/9 + 4y/9 - z/9)^2$
79. valores propios 1, 2, 3, definida positiva; $z^2 + 2(4x/5 - 3y/5)^2 + 3(3x/5 + 4y/5)^2$
80. valores propios 9, 18, -81, indefinida;
$9(-x/9 - 4y/9 - 8z/9)^2 + 18(4x/9 + 7y/9 - 4z/9)^2 + 81(-8x/9 + 4y/9 - z/9)^2$
81. valores propios 0, 13, 39, semidefinida positiva;
$13(12x/13 + 3y/13 - 4z/13)^2 + 39(4x/13 - 12y/13 + 3z/13)^2$
82. valores propios -13, -39, -169, definida negativa; $-13(-3x/13 - 4y/13 - 12z/13)^2 - 39(12x/13 + 3y/13 - 4z/13)^2 - 169(4x/13 - 12y/13 + 3z/13)^2$
83. valores propios 81, 162, 0, semidefinida positiva;
$81(-x/9 - 4y/9 - 8z/9)^2 + 162(4x/9 + 7y/9 - 4z/9)^2$
84. valores propios -5, -15, 30, indefinida;
$-5(-2x/15 - y/3 - 14z/15)^2 - 15(2x/3 + 2y/3 - z/3)^2 + 30(-11x/15 + 10y/15 - 2z/15)^2$
85. valores propios 7, 21, -49, indefinida;
$7(2x/7 - 3y/7 - 6z/7)^2 + 21(-6x/7 + 2y/7 - 3z/7)^2 - 49(3x/7 + 6y/7 - 2z/7)^2$

Respuestas a los ejercicios del capítulo 2,

Capítulo 2, Sección 1 (página 110)

1. $f(1, 0) = f(0, 1) = 1$, $f(1, 1) = 2$. $f(x, y) = 0$ sólo para $(x, y) = (0, 0)$. $f(x, y) = 1$ para los puntos del círculo unitario $x^2 + y^2 = 1$.
2. $f(2, 3) = 5$, $f(x, 1) = x + 1$, $f(1, y) = 1 + y$, $f(x^{-1}, y^{-1}) = \frac{1}{x} + \frac{1}{y}$. $f(x, k - x) = k$. f manda al 0 a los puntos de la recta $y = -x$.
3. $f(2, 5) = 29/2$, $f(x, y) = \frac{1}{2}(x^2 + y^2)$.
4. $f(x, y) = \frac{x^2(y + 1)}{y - 1}$. Dominio $= \{(x, y) | y \neq 1\}$.

5. $f(0, 0, 0) = 1$, $f(\pm 1, \pm 1, \pm 1) = e^{-3}$, $f(\{x^2 + y^2 + z^2\}) = e^{-1}$. Los valores de $f(x, y, z)$ tienden a 0 cuando $\|(x, y, z)\|$ tiende a infinito.
6. **a.** $\{(x, y)|x > y^2\}$; **b.** $\{(x, y)|x = y^2\}$; **c.** $\{(x, y)|x < y^2\}$
7. Dominio $= \mathbb{R}^2$, rango $= \{0\}$.
8. Dominio $= \{(x, y)|y > -x - 1\}$, rango $= \{-1, 0, 1\}$.
9. $\{(x, y)|y \geq -x\}$. Los puntos del plano xy que están por encima o coinciden con la recta $y = -x$.
10. $\{(x, y)|y > -x\}$. Los puntos del plano xy que están por encima de la recta $y = -x$.
11. $\{(x, y)|x \geq 0, y \geq 0\}$. Los puntos del plano xy que están en el primer cuadrante, incluyendo los ejes.
12. $\{(x, y)|x > 0, y > 0\}$. Los puntos del plano xy que están en el primer cuadrante sin incluír los ejes.
13. $\{(x, y)|x \leq 0, y \leq 0\} \cup \{(x, y)|x \geq 0, y \geq 0\}$. Los puntos del plano en el primer y tercer cuadrantes incluyendo los ejes.
14. $\{(x, y)|x \geq 0, y \geq 0\}$ (ver ejercicio 11).
15. $\{(x, y)|y \geq 0, x \geq -\sqrt{y}\}$. Los puntos del plano xy en el primer cuadrante junto con los puntos que están por encima de la semiparábola $x = -\sqrt{y}$, incluyendo la frontera.
16. $\{(x, y)|x^2 + y^2 < 1\}$. Los puntos del plano xy que están dentro del círculo unitario, sin incluir la frontera.
17. $\mathbb{R}^2$.
18. $\{(x, y)|x \neq 0, y \neq 0\}$. Todo el plano xy, excepto el origen.
19. $\{(x, y)|y \geq -x\}$. (Ver ejercicio 9).
20. $\{(x, y)|y > 0, x > -y\} \cup \{(x, y)|y < 0, -y - 1 < x < -y\}$. Los puntos del plano xy que están por encima de la recta $y = -x$ en el segundo cuadrante, junto con los puntos que están entre las rectas $y = -x$, $y = -x - 1$ en el tercero y cuarto cuadrantes, sin incluir las fronteras.
21. $\mathbb{R}^2$.
22. $\{(x, y)| -1 - x \leq y \leq 1 - x\}$. Los puntos del plano xy que están entre las rectas $y = -x - 1$, $y = -x + 1$, incluyéndolas.
23. $\{(x, y)| -1 - x^2 \leq y \leq 1 - x^2\}$. Los puntos del plano xy que están entre las parábolas $y = -1 - x^2$, $y = 1 - x^2$, incluyéndolas.
24. $\mathbb{R}^2$.
25. $\{(x, y)|y \geq -2x\}$. Los puntos del plano xy que están por encima de la recta $y = -2x$, incluyéndola.
26. $\mathbb{R}^2$.
27. $\{(x, y)|2k \leq x^2 + y^2 \leq 2k + 1, k = 0, 1, 2, \ldots\}$. Los puntos del plano xy que están en los anillos circulares limitados por los círculos con centro en el origen de radios $2k$ (por adentro) y $2k + 1$ (por afuera).
28. $\{(x, y)|y \geq 0, \frac{2k+1}{2}\pi \leq x \leq \frac{2k+3}{2}\pi, k \text{ impar}\} \cup \{(x, y)|y \leq 0, \frac{2k+1}{2}\pi \leq x \leq \frac{2k+3}{2}\pi, k \text{ par}\}$.
29. $\mathbb{R}^3$;
30. $\{(x, y, z)|x \geq 0, y \geq 0, z \geq 0\}$.
31. $\{(x, y, z)|z > 1\}$;
32. $\mathbb{R}^4$.
33. $\{(x_1, x_2, x_3, x_4)|x_1 + x_2 + x_3 + x_4 > 0\}$.
34. No.

35. Sí.
36. No.
37. No.
38. Sí.
39. Sí.
40. Sí.
41. $(f+g)(x,y)=x^2+y^2+1, \mathbb{R}^2$; $(fg)(x,y)=x^2+y^2+1, \mathbb{R}^2$; $\left(\frac{f}{g}\right)(x,y)=x^2+y^2+1, \mathbb{R}^2$.
42. $(f+g)(x,y)=2x, \mathbb{R}^2$; $(fg)(x,y)=x^2-y^2, \mathbb{R}^2$; $\left(\frac{f}{g}\right)(x,y)=\frac{x+y}{x-y}$, $\{(x,y)|x\neq y\}$.
43. $(f+g)(x,y)=\sqrt{1+x+y}+\sqrt{x}+\sqrt{y}$, $\{(x,y)|x\geq 0, y\geq 0\}$;
$(fg)(x,y)=\sqrt{1+x+y}(\sqrt{x}+\sqrt{y})$, $\{(x,y)|x\geq 0, y\geq 0\}$;
$\left(\frac{f}{g}\right)(x,y)=\frac{\sqrt{1+x+y}}{\sqrt{x}+\sqrt{y}}$, $\{(x,y)|x\geq 0, y\geq 0\ (0,0)\}$.
44. $(f+g)(x,y,z)=\operatorname{sen}(x+y+z)+2\cos(x+y+z), \mathbb{R}^3$;
$(fg)(x,y)=\operatorname{sen}(2x+2y+2z), \mathbb{R}^3$;
$\left(\frac{f}{g}\right)(x,y)=\frac{1}{2}\tan(x+y+z)$, $\{(x,y,z)|x+y+z\neq \frac{2k+1}{2}\pi, k\in\mathbb{Z}\}$.
45. $\{c\}$.
46. No.
47. Sí.

Capítulo 2, Sección 2 (página 123)

1. **a.** Dominio $=U$. Gráfica de g igual a gráfica de f, movida k unidades en el eje z; **b.** Dominio $=\{(x,y)|(x-x_0, y-y_0)\in U\}$. Gráfica de g igual a gráfica de f, movida x_0 unidades en el eje x y y_0 unidades en el eje y; **c.** Dominio $=\{(x,y)|(-x,-y)\in U\}$. Gráfica de g igual a gráfica de f, puesta simétricamente respecto del origen; **d.** Dominio $=U$. Gráfica de g es la reflexión de la gráfica de f en el plano xy.
2. **a.** $V=(3,-2,-1)$, $P=(0,0,3)$; **b.** $V=(-1,1,-1)$, $P=(0,0,-1)$; **c.** $V=(1,0,-1)$, $P=(0,0,0)$; **d.** $V=(0,1,-1)$, $P=(0,0,0)$; **e.** $V=(1,0,-1)$, $P=(0,0,7)$; **f.** $V=(0,0,-1)$, $P=(0,0,2)$.
3. **a.** $f(x)=x$; **b.** $f(x)=(x-2)^2$; **c.** $f(x)=\begin{cases} x & \text{si } x<1 \\ 1 & \text{si } 1\leq x\leq 3 \\ x-2 & \text{si } x>3\end{cases}$
d. $f(x)=[x]-x+1$; **e.** $f(x)=1$.
4. Verdadero.
5. **a.** $f(x,y)=x^2+y^2$, $c=0$; **b.** $f(x,y)=\operatorname{sgn}(\sqrt{x}\sqrt{y})$, $c=1$; **c.** $f(x,y)=k$, $c=k$.
6. Son líneas rectas paralelas al plano xy.
7. Son elipses con ecuación $ax^2+by^2=c$.
8. $\text{mín}(-1,1)=-1$, $\text{mín}(3,\pi)=3$, $\text{mín}(3,e)=e$.
9. $\text{máx}(2^{1/2},2^{3/2})=2^{3/2}$, $\text{máx}(2^{-1/2},2^{-3/2})=2^{-1/2}$, $z=\text{mín}(x,y)$.
10. **a.** $|x|=c$, si $|y|\geq|x|$; $|y|=c$, si $|y|<|x|$
b. $|x|=c$, si $|y|\leq|x|$; $|y|=c$, si $|y|>|x|$
c. $x^2=c$, si $y\geq x^2$; $y=c$, si $y<x^2$.
11. **a.** $f(x,y)=\operatorname{sen}x-y+1$; **b.** $f(x,y)=\sqrt{x^6+\ln^8 x}-y-7$; **c.** $f(x,y)=y^4x+x^3y+121$;
$f(x,y)=\operatorname{sgn}\left(\frac{2}{x^2+y^2+1}-1\right)-1$.

12. $f(x, y) = \phi(x) - y + c$; $U = I \times \mathbb{R}$.
16. $y = |x| - c$
17. $|y| = x - c$
18. $y = x - c, c \geq 0$
19. $y = \dfrac{c^2}{x}, c \geq 0$
20. $y = \dfrac{x}{c}$
21. $\left(x - \dfrac{1}{c}\right)^2 + y^2 = \left(\dfrac{1}{c}\right)^2$
22. $\left(y - \dfrac{1}{c}\right)^2 + x^2 = \left(\dfrac{1}{c}\right)^2$.
23. $y = \dfrac{c}{\operatorname{sgn} x}$
24. $y = \operatorname{sen} c - x, c \in \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$
25. **a.** $f(x, y, z) = x^2 + y^2 + z^2, c = 0$; **b.** $f(x, y, z) = (x - y)^2 + (x - z)^2, c = 0$;
c. $f(x, y, z) = \operatorname{sgn}(\sqrt{x}\sqrt{y}\sqrt{z}), c = 1$.
26. Planos paralelos, con vector normal (a, b, c).
27. **a.** $f(x, y, z) = x^2 + y^2 - z + 1$; **b.** $f(x, y, z) = \ln^2(\operatorname{sen}^4(x + y^8) + 7) - 7 - z$;
c. $f(x, y, z) = xz^3 + x^2y^5z^2 - 23yz + 254$; **d.** $f(x, y, z) = \operatorname{sgn}\left(\dfrac{2}{x^2 + y^2 + z^2 + 1} - 1\right) - 1$.
28. $f(x, y, z) = \phi(x, y) - z + c$, Dominio $= U \times \mathbb{R} \subseteq \mathbb{R}^3$.
31. $c = 0$, es el cono $z^2 = x^2 + y^2$; $c \neq 0$, son hiperboloides de una hoja.
32. $c = 0$, es el cono $y^2 = x^2 + z^2$; $c \neq 0$, son hiperboloides de una hoja.
33. $c = 0$, es el cono $x^2 = y^2 + z^2$; $c \neq 0$, son hiperboloides de dos hojas.
34. $c = 0$, es el origen; $c > 0$, son elipsoides.
35. $c = 0$, es el eje z menos el origen; $c \neq 0$ son paraboloides.

Capítulo 2, Sección 3 (página 139)

1. **a.** $\{x \in \mathbb{R} \mid |x - 3| < 0.5\}$; **b.** $\{(x, y) \in \mathbb{R}^2 \mid \|(x, y) - (2, -3)\| < 1\}$;
c. $\{(x, y, z) \in \mathbb{R}^3 \mid \|(x, y, z) - (1, 1, 4)\| < 2\}$;
d. $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid \|(x_1, x_2, x_3, x_4, x_5) - (2, -1, 9, 3, 5)\| < 1\}$.
5. Verdadero.
10. Abierto.
11. Abierto.
12. Abierto.
13. Cerrado.
14. Abierto.
15. Cerrado.
16. Abierto y cerrado (es el conjunto vacío).
17. Abierto.
18. Cerrado.
19. Abierto.
20. Ni abierto ni cerrado.
21. Abierto.
22. Cerrado.

23. Abierto y cerrado (es todo el espacio $\mathbb{R}^2$).
24. Abierto.
25. Abierto.
26. Abierto.
27. Abierto.
28. Abierto.
29. Cerrado.
30. Cerrado.
31. $\delta \leq 0.025$
32. $\delta \leq 0.013333$
33. $\delta \leq 0.0285714$
35. **a.** $\mathbb{R}^2 - \{(0,0)\}$.
36. **a.** $\mathbb{R}^2 - \{(0,0)\}$,
37. **a.** $\mathbb{R}^2 - \{(0,0)\}$.
38. **a.** $\mathbb{R}^2 - \{(0,0)\}$.
39. $\lim_{x\to 0} f(x,0) = 1 \neq -1 = \lim_{y\to 0} f(0,y)$.
40. $\lim_{x\to 0} f(x,0) = 0 \neq 1/2 = \lim_{x\to 0} f(x,x)$.
41. $\lim_{x\to 0} f(x,0) = 0 \neq 1/2 = \lim_{y\to 0} f(y^2,y)$.
42. $\lim_{x\to 0} f(x,0) = 0 \neq 2/7 = \lim_{x\to 0} f(x,x)$.
43. $\lim_{x\to 0} f(x,0) = 0 \neq 1/2 = \lim_{x\to 0} f(x^4,x^6)$.
44. $\lim_{x\to 0} f(x,0) = 0 \neq 1/2 = \lim_{x\to 0} f(x,x^4)$.
45. $\lim_{x\to 0} f(x,0) = 0 \neq 4 = \lim_{x\to 0} f(x^3,x^9)$.
46. $\lim_{x\to 0} f(x,0) = 0 \neq 1/2 = \lim_{y\to 0} f(y^3,y)$.
48. Dominio $= \{(x,y,z) | x+y-z \neq 0\}$; $\lim_{x\to 0} f(x,0,0) = 1$, $\lim_{z\to 0} f(0,0,z) = -1$.
49. Dominio $= \{(x,y,z) | |x| \neq |y|\}$; $\lim_{x\to 0} f(x,0,0) = 2$, $\lim_{y\to 0} f(0,y,0) = -1$.
50. Dominio $= \{(x,y,z) | x^3+y^3+z^3 \neq 0\}$; $\lim_{x\to 0} f(x,0,0) = 0$, $\lim_{x\to 0} f(x,x,x) = 1/3$.
51. Dominio $= \{(x,y,z) | x^2+z^2 \neq 0\}$; $\lim_{x\to 0} f(x,0,0) = 0$, $\lim_{x\to 0} f(x,x,x) = 1/2$.
54. $5/2$
55. 6
56. $3/2$
57. 2
58. 0
59. 6
60. $-4/7$
61. 2
63. Continua en $\mathbb{R}^2$
64. Continua en $\mathbb{R}^2 - \{(0,0)\}$
65. Continua en $\{(x,y) | y \neq \pm x\}$
66. Continua en $\mathbb{R}^2$
67. Continua en $\mathbb{R}^2$
68. Continua en $\{(x,y) | \operatorname{sen} y \neq 0\}$

69. Continua en $\mathbb{R}^2$
70. Continua en $\mathbb{R}^2 - \{(0,0)\}$
71. Continua en $\mathbb{R}^2 - \{(0,0)\}$
72. Continua en $\mathbb{R}^2$
73. No.
74. $f(0,0) = 0$
75. No.
76. $f(0,0) = 0$
77. No.

Capítulo 2, Sección 4 (página 152)

1. $\frac{\partial}{\partial x}x^4y^5 = 4x^3y^5.$

2. $\frac{\partial}{\partial x}(3y^2 \operatorname{sen} x^2 + \tan^2 x) = 6xy^2 \cos x^2 + 2 \tan x \sec^2 x.$

3. $\frac{\partial}{\partial y}\left(\ln \frac{y}{x} + 3 \ln \frac{x}{y}\right) = -\frac{2}{y}.$

4. $\frac{\partial}{\partial x}\sqrt{xy \operatorname{sen} z} = \frac{\sqrt{y \operatorname{sen} z}}{2\sqrt{x}}.$

5. $\frac{\partial}{\partial z}\left(\frac{xy + z^2}{zy} + z \cos^5 z^4\right)^{1/3} =$

$\frac{1}{3}\left(\frac{xy + z^2}{zy} + z \cos^5 z^4\right)^{-2/3}\left(\frac{z^2 - xy}{z^2 y} + \cos^5 z^4 - 20z^4 \cos^4 z^4 \operatorname{sen} z^4\right).$

6. $\frac{\partial}{\partial x}e^{(xyz)^2} = 2xy^2z^2e^{(xyz)^2}$

7. $\frac{\partial}{\partial z}(y \operatorname{sen} xz) = xy \cos xz.$

8. $\frac{\partial f}{\partial x} = 3(4x^2y^4 - 3x^2 + 8y^3)^2(8xy^4 - 6x)$; $\frac{\partial f}{\partial y} = 3(4x^2y^4 - 3x^2 + 8y^3)^2(16x^2y^3 + 24y^2)$

9. $\frac{\partial f}{\partial x} = -\frac{y}{x(x^2 - y^2)^{1/2}} - \frac{x}{y(x^2 - y^2)^{1/2}}$; $\frac{\partial f}{\partial y} = \frac{1}{(x^2 - y^2)^{1/2}} + \frac{x^2}{y^2(x^2 - y^2)^{1/2}}.$

10. $\frac{\partial f}{\partial x} = \frac{2y(x-1)}{(x-y)^2}$; $\frac{\partial f}{\partial y} = \frac{2x(y-1)}{(x-y)^2}.$

11. $\frac{\partial f}{\partial x} = \frac{x^2y + x^2y^2 + x^2 - y^2}{x^2y}$; $\frac{\partial f}{\partial y} = \frac{y^2x + x^2y^2 - x^2 + y^2}{y^2x}.$

12. $\frac{\partial f}{\partial x} = yx^{y-1} + y^x \ln y$; $\frac{\partial f}{\partial y} = x^y \ln x + xy^{x-1}.$

13. $\frac{\partial f}{\partial x} = x^x(\ln x + 1) + yx^{y-1}y^x + x^y y^x \ln y$; $\frac{\partial f}{\partial y} = y^y(\ln y + 1) + x^y xy^{x-1} + y^x x^y \ln x.$

14. $\frac{\partial f}{\partial x} = \left[\ln(2x + 3y) + \frac{2x}{2x + 3y}\right](2x + 3y)^x + 2y(2x + 3y)^{y-1}$;

$\frac{\partial f}{\partial y} = 3x(2x + 3y)^{x-1} + \left[\ln(2x + 3y) + \frac{3y}{2x + 3y}\right](2x + 3y)^y.$

15. $\frac{\partial f}{\partial x} = \frac{yy^{x^y}x^y \ln y}{x} + x^{y^x}y^x\left[\ln y \ln x + \frac{1}{x}\right] + (x^y)^x(y^x)^y(\ln x^y + y \ln y + y)$;

$\frac{\partial f}{\partial y} = y^{x^y}x^y\left[\ln y \ln x + \frac{1}{y}\right] + \frac{xx^{y^x}y^x \ln x}{y} + (x^y)^x(y^x)^y(\ln y^x + x \ln x + x).$

16. $\dfrac{\partial f}{\partial x} = -\dfrac{4x}{(x^2+y^2+1)\ln^3(x^2+y^2+1)}$; $\dfrac{\partial f}{\partial y} = -\dfrac{4y}{(x^2+y^2+1)\ln^3(x^2+y^2+1)}$.

17. $\dfrac{\partial f}{\partial x} = (2y)^x \ln(2y)$; $\dfrac{\partial f}{\partial y} = \dfrac{x(2y)^x}{y} + 2^y \ln 2$.

18. $\dfrac{\partial f}{\partial x} = \ln y - \dfrac{y}{x}$; $\dfrac{\partial f}{\partial y} = \dfrac{x}{y} - \ln x$.

19. $\dfrac{\partial f}{\partial x} = -\dfrac{x}{\sqrt{1-x^2-y^2}\sqrt{x^2+y^2}}$; $\dfrac{\partial f}{\partial y} = -\dfrac{y}{\sqrt{1-x^2-y^2}\sqrt{x^2+y^2}}$.

20. $\dfrac{\partial f}{\partial x} = \dfrac{y(3\cos x+2)(3\operatorname{sen} x+2x)^y}{(3\operatorname{sen} x+2x)^{2y+1}+3\operatorname{sen} x+2x}$; $\dfrac{\partial f}{\partial y} = \dfrac{\ln(3\operatorname{sen} x+2x)(3\operatorname{sen} x+2x)^y}{(3\operatorname{sen} x+2x)^{2y}+1}$.

21. $\dfrac{\partial f}{\partial x} = \dfrac{2x}{\sqrt{x^2+y^2}(x^2+y^2-1)}$; $\dfrac{\partial f}{\partial y} = \dfrac{2y}{\sqrt{x^2+y^2}(x^2+y^2-1)}$.

22. $\dfrac{\partial f}{\partial x} = \dfrac{2x}{\sqrt{x^2+y^2+z^2}(x^2+y^2+z^2-1)}$; $\dfrac{\partial f}{\partial y} = \dfrac{2y}{\sqrt{x^2+y^2+z^2}(x^2+y^2+z^2-1)}$;
$\dfrac{\partial f}{\partial z} = \dfrac{2z}{\sqrt{x^2+y^2+z^2}(x^2+y^2+z^2-1)}$.

23. $\dfrac{\partial f}{\partial x} = yx^{y-1} + zx^{z-1} + y^x \ln y + z^x \ln z$; $\dfrac{\partial f}{\partial y} = x^y \ln x + xy^{x-1} + zy^{z-1} + z^y \ln z$;
$\dfrac{\partial f}{\partial z} = x^z \ln x + y^z \ln y + xz^{x-1} + yz^{y-1}$.

24. $\dfrac{\partial f}{\partial x} = \dfrac{2yzx^{yz}}{x} + 2zy^{xz}\ln y + 2yz^{xy}\ln z$;
$\dfrac{\partial f}{\partial y} = \dfrac{2xzy^{xz}}{y} + 2zx^{yz}\ln x + 2xz^{xy}\ln z$;
$\dfrac{\partial f}{\partial z} = \dfrac{2xyz^{xy}}{z} + 2yx^{yz}\ln x + 2xy^{xz}\ln y$.

25. $\dfrac{\partial f}{\partial x} = 2x\arctan\sqrt{1+y+z}$; $\dfrac{\partial f}{\partial y} = \dfrac{x^2\sqrt{1+y+z}+2(2+y+z)}{2(2+y+z)(1+y+z)}$;
$\dfrac{\partial f}{\partial z} = \dfrac{x^2\sqrt{1+y+z}+2(2+y+z)}{2(2+y+z)(1+y+z)}$.

26. $\dfrac{\partial f}{\partial x} = yz\tan^2 x + z\cos y + y\operatorname{sen} z + yz$. $\dfrac{\partial f}{\partial y} = z\tan x - xz\operatorname{sen} y + x\operatorname{sen} z$;
$\dfrac{\partial f}{\partial z} = y\tan x + x\cos y + xy\cos z$.

27. $\dfrac{\partial f}{\partial x} = 2x^2y^3z^4\tan^4 z\cos^3 y\operatorname{sen} x\cos x + 2xy^3z^4\tan^4 z\cos^3 y\operatorname{sen}^2 x$;
$\dfrac{\partial f}{\partial y} = \operatorname{sen}^2 x(3x^2y^2z^4\tan^4 z\cos^3 y - 3x^2y^3z^4\tan^4 z\operatorname{sen} y\cos^2 y)$;
$\dfrac{\partial f}{\partial z} = \cos^3 y\operatorname{sen}^2 x(4x^2y^3z^4\tan^3 z + 4x^2y^3z^3\tan^4 z + 4x^2y^3z^4\tan^5 z)$.

28. $\dfrac{\partial f}{\partial x} = \dfrac{(y+z)(x^2-yz)}{x^3(y+z)+x^2(y^2+z^2)+xyz(y+z)}$;
$\dfrac{\partial f}{\partial y} = \dfrac{(y^2-xz)(x+z)}{y(x^2(y+z)+x(y^2+z^2)+yz(y+z))}$;
$\dfrac{\partial f}{\partial z} = \dfrac{(z^2-xy)(x+y)}{z(x^2(y+z)+x(y^2+z^2)+yz(y+z))}$.

29. $\dfrac{\partial f}{\partial x} = x^{y+z+u}z^{x+y+u}\left(\ln z + \dfrac{y+z+u}{x}\right)$;

$\dfrac{\partial f}{\partial y} = x^{y+z+u}z^{x+y+u}\left(\ln x + \ln z\right)$;

$\dfrac{\partial f}{\partial z} = x^{y+z+u}z^{x+y+u}\left(\ln z + \dfrac{x+y+u}{z}\right)$

30. $\dfrac{\partial f}{\partial x_1} = \dfrac{1}{x_2}$; $\dfrac{\partial f}{\partial x_i} = \dfrac{1}{x_{i+1}} - \dfrac{x_{i-1}}{(x_i)^2}$, $i = 2, \ldots, n-1$; $\dfrac{\partial f}{\partial x_n} = -\dfrac{x_{n-1}}{(x_n)^2}$.

31. $\dfrac{\partial f}{\partial x}(1,1,1) = \dfrac{\partial f}{\partial y}(1,1,1) = \dfrac{\partial f}{\partial z}(1,1,1) = 2$.

32. $\dfrac{\partial f}{\partial x}(1,1,1) = \dfrac{\partial f}{\partial y}(1,1,1) = \dfrac{\partial f}{\partial z}(1,1,1) = 2$.

33. $\displaystyle\sum_{i=1}^{n} \dfrac{\partial f}{\partial x_i}(1, 1, \ldots, 1) = n$.

38. **a.** 18; **b.** 12.

41. **a.** $\{(x, y) | y > x\}$; **b.** $\{(x, y) | y < x\}$; **c.** La recta $y = x$; **d.** $\dfrac{\partial f}{\partial x} = -g(x)$, $\dfrac{\partial f}{\partial y} = g(y)$.

42. **a.** $\{(x, y) | |y| > |x|\}$; **b.** $\{(x, y) | |y| < |x|\}$; **c.** Las rectas $y = \pm x$; **d.** $\dfrac{\partial f}{\partial x} = -g(x)$, $\dfrac{\partial f}{\partial y} = g(y)$.

43. $\dfrac{\partial f}{\partial x} = yg(xy) - g(x)$; $\dfrac{\partial f}{\partial y} = xg(xy)$.

44. $\dfrac{\partial f}{\partial x} = g(x-y) - g(x+y)$; $\dfrac{\partial f}{\partial y} = -g(x+y) - g(x-y)$.

45. $\dfrac{\partial f}{\partial x} = (x^2+y^2)\left[-yg(xy)\right] + 2x\displaystyle\int_{xy}^{y} g(t)dt$; $\dfrac{\partial f}{\partial y} = (x^2+y^2)\left[-xg(xy) + g(y)\right] + 2y\displaystyle\int_{xy}^{y} g(t)dt$.

46. $\dfrac{\partial f}{\partial x} = -g(x^y)yx^{y-1} + g(y^x)y^x \ln y$; $\dfrac{\partial f}{\partial y} = -g(x^y)x^y \ln x + g(y^x)xy^{x-1}$.

47. $\dfrac{\partial f}{\partial x} = -g\left(\displaystyle\int_x^y g(t)dt\right)g(x)$; $\dfrac{\partial f}{\partial y} = g(y)$.

48. $\dfrac{\partial f}{\partial x} = g\left(\displaystyle\int_x^y g(t)dt\right)g(x) + g\left(\displaystyle\int_y^x g(t)dt\right)g(x)$;

$\dfrac{\partial f}{\partial y} = -g\left(\displaystyle\int_x^y g(t)dt\right)g(y) - g\left(\displaystyle\int_y^x g(t)dt\right)g(y)$.

49. $\dfrac{\partial f}{\partial x} = g(x+y+z) - g(xyz)yz$; $\dfrac{\partial f}{\partial y} = g(x+y+z) - g(xyz)xz$;

$\dfrac{\partial f}{\partial z} = g(x+y+z) - g(xyz)xy$.

50. $\dfrac{\partial f}{\partial x} = g\left(\displaystyle\int_{x+y+z}^{\int_x^y g(t)dt} g(t)dt\right)\left[-g(x)g\left(\displaystyle\int_x^y g(t)dt\right) - g(x+y+z)\right] - g(x+y+z)$;

$\dfrac{\partial f}{\partial y} = g\left(\displaystyle\int_{x+y+z}^{\int_x^y g(t)dt} g(t)dt\right)\left[g(y)g\left(\displaystyle\int_x^y g(t)dt\right) - g(x+y+z)\right] - g(x+y+z)$;

$\dfrac{\partial f}{\partial z} = g\left(\displaystyle\int_{x+y+z}^{\int_x^y g(t)dt} g(t)dt\right)\left[-g(x+y+z)\right] - g(x+y+z)$.

51. $\dfrac{\partial f}{\partial x} = g'(x)$; $\dfrac{\partial f}{\partial y} = 5h'(y)$.

52. $\dfrac{\partial f}{\partial x} = 2g'(x)\Big(h(y) + g(x)\Big)$; $\dfrac{\partial f}{\partial y} = 2g(x)h'(y) + 2yh'(y^2)$.

53. $\dfrac{\partial f}{\partial x} = \dfrac{h'(x)}{1 + g^2(y)}$; $\dfrac{\partial f}{\partial y} = -\dfrac{1 + h(x)}{(1 + g^2(y))^2}\Big(2g(y)g'(y)\Big)$.

54. $\dfrac{\partial f}{\partial x} = g(h(y))h'(x) + g(y)h'(g(x))g'(x)$; $\dfrac{\partial f}{\partial y} = h(x)g'(h(y))h'(y) + h(g(x))g'(y)$.

55. $\dfrac{\partial f}{\partial x} = \operatorname{sen} h(g(y))g'(h(x))h'(x) - h(g(y)) \operatorname{sen} g(h(x))g'(h(x))h'(x)$;

$\dfrac{\partial f}{\partial y} = g(h(x)) \cos h(g(y))h'(g(y))g'(y) + \cos g(h(x))h'(g(y))g'(y)$.

56. $\dfrac{\partial f}{\partial x} = \dfrac{2g(x)g'(x)}{1 + g^2(x) + g^4(y) + g^6(z)}$; $\dfrac{\partial f}{\partial y} = \dfrac{4g^3(x)g'(x)}{1 + g^2(x) + g^4(y) + g^6(z)}$;

$\dfrac{\partial f}{\partial z} = \dfrac{6g^5(z)g'(z)}{1 + g^2(x) + g^4(y) + g^6(z)}$.

57. $\dfrac{\partial f}{\partial x} = g(z)g'(g(x)g(y))g(y)g'(x)$; $\dfrac{\partial f}{\partial y} = g(z)g'(g(x)g(y))g(x)g'(y)$;

$\dfrac{\partial f}{\partial z} = g(g(x)g(y))g'(z)$.

58. $\dfrac{\partial f}{\partial x} = g'(g(x)g(g(y)g(z)))g(g(y)g(z))g'(x)$;

$\dfrac{\partial f}{\partial y} = g'(g(x)g(g(y)g(z)))g'(g(y)g(z))g(x)g(z)g'(y)$;

$\dfrac{\partial f}{\partial z} = g'(g(x)g(g(y)g(z)))g'(g(y)g(z))g(x)g(y)g'(z)$.

59. $\dfrac{\partial f}{\partial x} = h(y)(g(x))^{h(y)-1}g'(x)$

$\dfrac{\partial f}{\partial y} = (g(x))^{h(y)}(\ln g(x))h'(y) + g(z)(h(y))^{g(z)-1}h'(y)$

$\dfrac{\partial f}{\partial z} = h(y)^{g(z)}(\ln h(y))g'(z)$

60. $\dfrac{\partial f}{\partial x} = (h(y))^{g(z)}(g(x))^{(h(y))^{g(z)}-1}g'(x)$;

$\dfrac{\partial f}{\partial y} = (g(x))^{(h(y))^{g(z)}}(\ln g(x))g(z)(h(y))^{g(z)-1}h'(y)$;

$\dfrac{\partial f}{\partial z} = (g(x))^{(h(y))^{g(z)}}(\ln(g(x))^{h(y)})g'(z)$

61. $\dfrac{\partial f}{\partial x} = (h(g(y)))^{g(h(z))}(g(h(x)))^{(h(g(y)))^{g(h(z))}-1}g'(h(x))h'(x)$;

$\dfrac{\partial f}{\partial y} = (g(h(x)))^{(h(g(y)))^{g(h(z))}}(\ln g(h(x)))g(h(z))(h(g(y)))^{g(h(z))-1}h'(g(y))g'(y)$;

$\dfrac{\partial f}{\partial z} = (g(h(x)))^{(h(g(y)))^{g(h(z))}}(\ln(g(h(x)))^{h(g(y))})g'(h(z))h'(z)$

73. $\dfrac{\partial F}{\partial x} = yg'(xy)$; $\dfrac{\partial F}{\partial y} = xg'(xy)$

74. $\dfrac{\partial F}{\partial x} = 6xg'(3x^2 + 7y^2)$; $\dfrac{\partial F}{\partial y} = 14yg'(3x^2 + 7y^2)$

75. $\dfrac{\partial F}{\partial x} = \dfrac{\partial F}{\partial y} = 2g(x + y)g'(x + y)$

76. $\dfrac{\partial F}{\partial x} = g'(x+y^2) + 2xg'(x^2+y)$; $\dfrac{\partial F}{\partial y} = 2yg'(x+y^2) + g'(x^2+y)$

77. $\dfrac{\partial F}{\partial x} = 9x^2g(3x^3y^4)g'(3x^3+y^4) + 9x^2y^4g(3x^3+y^4)g'(3x^3y^4)$;

$\dfrac{\partial F}{\partial y} = 4y^3g(3x^3y^4)g'(3x^3+y^4) + 12x^3y^3g(3x^3+y^4)g'(3x^3y^4)$

78. $\dfrac{\partial F}{\partial x} = \dfrac{4ag^3(ax+by+c)}{4+g^4(ax+by+c)}$; $\dfrac{\partial F}{\partial y} = \dfrac{4bg^3(ax+by+c)}{4+g^4(ax+by+c)}$

79. $\dfrac{\partial F}{\partial x} = \dfrac{(\cos g(x\cos y + y\,\mathrm{sen}\,x))g'(x\cos y + y\,\mathrm{sen}\,x)(\cos y + y\cos x)}{1+(1+\mathrm{sen}\,g(x\cos y + y\,\mathrm{sen}\,x))^2}$

$\dfrac{\partial F}{\partial y} = \dfrac{(\cos g(x\cos y + y\,\mathrm{sen}\,x))g'(x\cos y + y\,\mathrm{sen}\,x)(-x\,\mathrm{sen}\,y + \mathrm{sen}\,x)}{1+(1+\mathrm{sen}\,g(x\cos y + y\,\mathrm{sen}\,x))^2}$

80. $\dfrac{\partial F}{\partial x} = 7g^6(g^6(x^5+y^4)g^3(x^2+y))g'(g^6(x^5+y^4)g^3(x^2+y)) *$

$* \left[6xg^6(x^5+y^4)g^2(x^2+y)g'(x^2+y) + 30x^4g^5(x^5+y^4)g'(x^5+y^4)g^3(x^2+y)\right]$;

$\dfrac{\partial F}{\partial y} = 7g^6(g^6(x^5+y^4)g^3(x^2+y))g'(g^6(x^5+y^4)g^3(x^2+y)) *$

$* \left[3g^6(x^5+y^4)g^2(x^2+y)g'(x^2+y) + 24y^3g^5(x^5+y^4)g'(x^5+y^4)g^3(x^2+y)\right]$

Capítulo 2, Sección 5 (página 165)

1. $\dfrac{\partial}{\partial v}(xy^2) = \dfrac{2xy+y^2}{\sqrt{2}}$, $\mathbf{v} = \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$.

2. $\dfrac{\partial}{\partial v}(\sqrt{xy}) = \dfrac{x-\sqrt{3}y}{4\sqrt{xy}}$, $\mathbf{v} = \left(\dfrac{\sqrt{3}}{2}, \dfrac{1}{2}\right)$.

3. $\dfrac{\partial}{\partial x}(\ln \mathrm{sen}^2(x^4y)) = 8x^3y\cot(x^4y)$.

4. $\dfrac{\partial}{\partial y}(y^2\cos^3(xy)) = 2y\cos^3(xy) - 3xy^2\cos^2(xy)\,\mathrm{sen}(xy)$.

5. $\dfrac{\partial}{\partial v}(xyz) = \dfrac{2yz+2xz-xy}{3}$, $\mathbf{v} = \left(\dfrac{2}{3}, \dfrac{2}{3}, -\dfrac{1}{3}\right)$.

6. $\dfrac{\partial}{\partial v}(x^2yz) = \dfrac{-\sqrt{3}x^2z - x^2y}{2}$, $\mathbf{v} = \left(0, -\dfrac{\sqrt{3}}{2}, -\dfrac{1}{2}\right)$.

7. $\dfrac{\partial f}{\partial v} = \dfrac{1}{\sqrt{2}}$.

8. $\dfrac{\partial f}{\partial v} = -\dfrac{1}{\sqrt{2}}$.

9. $\dfrac{\partial f}{\partial v}(0,0) = 0$.

10. $\dfrac{\partial f}{\partial v} = y^2 + 2xy$.

11. $\dfrac{\partial f}{\partial v} = 0$.

12. $\dfrac{\partial f}{\partial v} = \mathrm{sen}\,y^3\cos(x^5 + \tan y^3)$

13. $\dfrac{\partial f}{\partial v} = \dfrac{yz-2xz-2xy}{3}$.

14. $\dfrac{\partial f}{\partial v} = \dfrac{yzu-2xzu-2xyu}{3}$,

15. $\dfrac{\partial f}{\partial v} = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n$

16. $\dfrac{\partial f}{\partial v}(p) = \lim\limits_{t\to 0} \dfrac{f(p+tv)-f(p)}{t} = \lim\limits_{t\to 0} \dfrac{f(p+tu/\|u\|)-f(p)}{t} = \overset{s=t/\|u\|}{\longrightarrow} =$
$\lim\limits_{s\to 0} \dfrac{f(p+su)-f(p)}{\|u\|s} = \dfrac{1}{\|u\|}\dfrac{\partial f}{\partial u}(p)$

17. $\dfrac{\partial f}{\partial u} = -\dfrac{22}{\sqrt{38}}$.

Capítulo 2, Sección 6 (página 177)

1. $r(h_1, h_2) = 0$.
2. $r(h_1, h_2) = 0$.
3. $r(h_1, h_2) = 0$.
4. $r(h_1, h_2) = 0$.
5. $r(h_1, h_2) = 3h_1^2 + 9h_2^2$.
6. $r(h_1, h_2) = 4(3h_2^2 + h_2^3 + 6h_1h_2 + 6h_1h_2^2 + 2h_1h_2^3 + h_1^2 + 3h_1^2h_2 + 3h_1^2h_2^2 + h_1^2h_2^3)$
7. $r(h_1, h_2) = h_1 \operatorname{sen} h_2$.
8. $r(h_1, h_2, h_3) = 3h_1^2 + h_1^3 + 3h_2^2 + h_3^2$.
9. $r(h_1, h_2, h_3) = h_1((h_2+1)^2(h_3+1)^3 - 1) + 2(h_2^2(h_3+1)^3 + 2h_2((h_3+1)^3 - 1) + h_3^3 + 3h_3^2)$.
10. $r(h_1, h_2, h_3) = \exp(h_1 + h_2 + h_3) - h_1 - h_2 - h_3 - 1$.
16. Dominio $= \mathbb{R}^2$, diferenciable.
17. Dominio $= \mathbb{R}^2$, diferenciable.
18. Dominio $= \mathbb{R}^2$, diferenciable.
19. Dominio $= \mathbb{R}^2$, diferenciable.
20. Dominio $= \mathbb{R}^2$, diferenciable.
27.

	a	b	c	d	e	f	g	h
a	V	V	V	V	V	V	V	F
b	F	V	F	V	F	F	F	F
c	F	F	V	V	F	F	F	F
d	F	F	F	V	F	F	F	F
e	F	V	V	V	V	F	F	F
f	F	V	V	V	F	V	F	F
g	F	V	V	V	V	V	V	F
h	V	V	V	V	V	V	V	V

Capítulo 2, Sección 7 (página 191)

2. $\dfrac{\partial f}{\partial u}(x_0, y_0) = \dfrac{1}{\sqrt{13}}(2a + 3b)$.
3. $\dfrac{\partial f}{\partial u}(0, 0) = 0$.
4. $\dfrac{\partial f}{\partial u}(0, 0) = 0$.
5. $\dfrac{\partial f}{\partial u}(0, 0) = 0$.
6. $\dfrac{\partial f}{\partial u}(1, \pi) = -\dfrac{2}{\sqrt{13}}$.
7. $\dfrac{\partial f}{\partial u}(1, 1, 1) = 0$.

8. $\frac{\partial f}{\partial u}(1, 2, 0) = 0.$
9. $\frac{\partial f}{\partial u}(x_0, y_0, z_0) = \sqrt{a^2 + b^2 + c^2}.$
10. $\frac{\partial f}{\partial u}(1, 1, 1) = \sqrt{3}.$
11. **a.** $-\frac{25}{\sqrt{13}}$; **b.** $-\frac{25}{\sqrt{2}}$; **c.** $\frac{5}{\sqrt{2}}$; **d.** $\frac{25}{\sqrt{2}}$.
12. $2\sqrt{3}$.
14. En la dirección de $\mathbf{v} = \left(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}}\right)$; En la dirección tangente al círculo unitario.
29. **a.** Son elipses.

Capítulo 2, Sección 8 (página 197)

1. **a.** $(6 - \operatorname{sen} 1, 3 - \operatorname{sen} 1)$; **b.** $(4, 4\ln 2)$; **c.** $(0, 1)$; **d.** $(1/2, -1/2, 1)$.
2. **a.** $\theta = \arccos \frac{20}{\sqrt{481}}$; **b.** $\theta = \arccos \frac{6}{7}$.
3. **a.** $\theta = \pi/2$; **b.** $\theta = \arccos \frac{196}{\sqrt{161504}}$.
4. $-6\mathbf{u} + (10, -10)$.
5. $5\mathbf{u} + (5, 0)$.
6. $\frac{2\pi}{13}\mathbf{u} + \left(\frac{9\pi}{13}, -\frac{6\pi}{13}\right)$
7. $0\mathbf{u} + (0, 0)$.
8. $\mathbf{u} + (0, 0, 0)$.
9. $0\mathbf{u} + (2, 2, 2)$.
10. $\frac{6}{19}\mathbf{u} + \left(\frac{2}{19}, \frac{4}{19}, -\frac{24}{19}\right)$
11. **a.** 4; **b.** 0.
12. **a.** En la dirección del vector (a, b), en donde $3a + 4b = 2$. **b.** En la dirección del vector (a, b), en donde $3a + 4b = 0$.

Capítulo 2, Sección 9 (página 205)

1. $(0, 0, 1)$
2. $(0, 0, 1)$
3. $(-1, 0, 1)$
4. $(-1, 0, 1)$
5. $(-y_0, -x_0, 1)$
6. $(-1, -1, 1)$
7. $(0, -e, 1)$
8. $(-8, -15, 1)$
9. $(-1, -1, 2)$
10. $(-1, 0, 1)$
11. $(1, 1, 1)$
12. $(1, 1, 1)$
13. $(3, 4, -5)$ en ambos puntos
14. $(1, 1, 1)$
15. $(1, 3, 2)$
20. $(0, 0, 0)$

21. en cualquier punto de la gráfica de $z = f(x, y)$
22. no existe punto alguno
23. $(8/31, -1/93, 35/279)$
24. no existe punto alguno
25. en los puntos del conjunto $\{(x, y, f(x, y)) | x^2 + y^2 = k^2\pi^2, k \neq 0\}$
26. $\frac{2}{\sqrt{3}}(1, 1, 1)$
27. $(-k, 3k/4, k)$ en donde $k = \pm\sqrt{\frac{8}{41}}$
28. no existe punto alguno;
29. $(2, 2, -1)$ y $(-2, 0, 1/5)$
30. en los puntos $(1 \pm 5\sqrt{2}/6, 2 \pm 2\sqrt{2}/3, -1 \pm 7\sqrt{2}/6)$

Capítulo 2, Sección 10 (página 214)

1. $z = 3x + 8y - 10$
2. $z = 0$
3. $z = x + (2\ln 2)(y - 1)$
4. $z = y$
5. $z = 0$
9. $-2x - 2y + 5z = 1$
10. $x = 0$
11. $x + 2y + 5z = 20$
12. $3x + 8y - 5z = 73/20$
14. $z = -3, z = 5/3$
16. $y + z = 1$
17. $56x - 112z + 3 = 0$
18. $-4x + 2y - z = 1$
19. $x + y - z = \pm\sqrt{345}/15$
21. $13x + 14y - 25z = -48 \pm 3\sqrt{110}$
22. $13x + 14y - 25z = 66 \pm 18\sqrt{3}$

Ejercicio	Planos tangentes paralelos al plano $x = 0$ en los puntos	Planos tangentes paralelos al plano $y = 0$ en los puntos	Planos tangentes paralelos al plano $z = 0$ en los puntos
23	$(\pm\sqrt{12}, 0, 0)$	$(0, \pm\sqrt{12/5}, 0)$	$(0, 0, \pm\sqrt{6/5})$
24	$(2 \pm \sqrt{12}, 3, -1)$	$(2, 3 \pm \sqrt{12/5}, -1)$	$(2, 3, -1 \pm \sqrt{6/5})$
25	$(-1, 1, 1)$ $(-3, 1, 1)$	$(-2, 1 \pm \sqrt{3}/3, 1)$	$(-2, 1, 0)$ $(-2, 1, 2)$
26	$(1 \pm \sqrt{26}/3,$ $2 \pm 5\sqrt{26}/39,$ $-1 \pm 2\sqrt{26}/39)$	$(1 \pm 2\sqrt{35}/21,$ $2 \pm \sqrt{35}/3,$ $-1 \pm 2\sqrt{35}/15)$	$(1 \pm 2\sqrt{38}/57,$ $2 \pm 7\sqrt{38}/57,$ $-1 \pm \sqrt{38}/3)$

27. $(\pm 4\sqrt{222}/37, \mp 3\sqrt{222}/37, \pm\sqrt{222}/37)$
28. $3x + 2y + 2z = \pm 34\sqrt{30}/15$
29. $z = 0, 2x + 2y - 2z = 1$

30. $3125\sqrt{2}/9$ unidades cúbicas

31. En los 8 puntos $\left(\pm\dfrac{a^2}{\sqrt{a^2+b^2+c^2}}, \pm\dfrac{b^2}{\sqrt{a^2+b^2+c^2}}, \pm\dfrac{c^2}{\sqrt{a^2+b^2+c^2}}\right)$

35. En el punto (2, −6, 3). El otro plano tangente es $2x - 6y + 3z + 49 = 0$

36. $x + 4y + 5z = -64 \pm 14\sqrt{66}$

37. $3x + 7y + 6z = 59$, $3x + 7y + 6z = -35$

38. $3x + 4y + 5z = 0$, $3x + 4y + 5z = 100$

39. $3x^2 + 3y^2 + 3z^2 + 2x - 4y - 4z - 13 = 0$

40. $ax + by + cz = ax_0 + by_0 + cz_0 \pm r\sqrt{a^2+b^2+c^2}$

41. **a.** $y = 0$, $y = 4z/3$. **b.** $x = 0$, $x = 4z/3$. **c.** $x = 0$, $y = 0$

Capítulo 2, Sección 11 (página 221)

1. $df(x) = 6x \operatorname{sen}^2 x^2 \cos x^2 dx$

2. $df(x, y) = \tan y dx + x \sec^2 y dy$

3. $df(x, y, z) = adx + bdy + cdz$

4. $df(x, y, z, u) = \cos x dx - \operatorname{sen} y dy + \dfrac{dz}{\sqrt{1-z^2}} - \dfrac{du}{\sqrt{1-u^2}}$

5. $df(x, y, z, u, w) = (yz+zw)dx+(xz+uw)dy+(xy+xw+uw)dz+(yw+zw)du+(yu+zu)dw$

8. $a_i h_i$

15. $x_0^2 + y_0^2$

Capítulo 2, Sección 12 (página 235)

1. $\dfrac{\partial^2 f}{\partial x^2} = -y \operatorname{sen} x$, $\dfrac{\partial^2 f}{\partial x \partial y} = \cos x + \cos y$, $\dfrac{\partial^2 f}{\partial y^2} = -x \operatorname{sen} y$

2. $\dfrac{\partial^2 f}{\partial x^2} = x^{y-2}(y^2 - y)$, $\dfrac{\partial^2 f}{\partial x \partial y} = x^{y-1}(y \ln x + 1)$, $\dfrac{\partial^2 f}{\partial y^2} = x^y \ln^2 x$

3. $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{-2xy^3}{(1+x^2y^2)^2}$, $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{1 - x^2y^2}{(1+x^2y^2)^2}$, $\dfrac{\partial^2 f}{\partial y^2} = \dfrac{-2x^3y}{(1+x^2y^2)^2}$

4. $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y^2} = -\dfrac{1}{(x+y)^2}$

5. $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial^2 f}{\partial y^2} = -\operatorname{sen} y \operatorname{sen} x$, $\dfrac{\partial^2 f}{\partial x \partial y} = \cos y \cos x$

6. Todas las parciales de segundo orden son iguales a cero

7. **a.** $\dfrac{\partial^2 F}{\partial x_i \partial x_j} = \begin{cases} f_i''(x_i) & \text{si } j = i \\ 0 & \text{si } j \neq i \end{cases}$

b. $\dfrac{\partial^2 F}{\partial x_i \partial x_j} = \begin{cases} 2(f_i(x_i)f_i''(x_i) + (f_i'(x_i))^2) & \text{si } j = i \\ 0 & \text{si } j \neq i \end{cases}$

c. $\dfrac{\partial^2 F}{\partial x_i \partial x_j} = \begin{cases} i f_i^{i-1}(x_i)f_i''(x_i) + i(i-1)f_i^{i-2}(x_i)(f_i'(x_i))^2 & \text{si } j = i \\ 0 & \text{si } j \neq i \end{cases}$

d. $\dfrac{\partial^2 F}{\partial x_i^2} = f_i''(x_i) \displaystyle\prod_{\substack{k=1 \\ k\neq i}}^{n} f_k(x_k)$, $\dfrac{\partial^2 F}{\partial x_i \partial x_j} = f_i'(x_i) f_j'(x_j) \displaystyle\prod_{\substack{k=1 \\ k\neq i,j}}^{n} f_k(x_k)$

e. $\dfrac{\partial^2 F}{\partial x_i^2} = 2(f_i(x_i)f_i''(x_i) + (f_i'(x_i))^2) \displaystyle\prod_{\substack{k=1 \\ k\neq i}}^{n} f_k^2(x_k)$,

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = 4 f_i(x_i) f_i'(x_i) f_j(x_j) f_j'(x_j) \prod_{\substack{k=1\\k\neq i\,j}}^{n} f_k^2(x_k)$$

f. $\frac{\partial^2 F}{\partial x_i^2} = (i f_i^{i-1}(x_i) f_i''(x_i) + i(i-1) f_i^{i-2}(x_i)(f_i'(x_i))^2) \prod_{\substack{k=1\\k\neq i}}^{n} f_k^k(x_k)$;

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = i j f_i^{i-1}(x_i) f_j^{j-1}(x_j) f_i'(x_i) f_j'(x_j) \prod_{\substack{k=1\\k\neq i\,j}}^{n} f_k^k(x_k)$$

18. $\frac{\partial^2 F}{\partial x^2} = 4g''(2x+3y)$, $\frac{\partial^2 F}{\partial x \partial y} = 6g''(2x+3y)$, $\frac{\partial^2 F}{\partial y^2} = 9g''(2x+3y)$

19. $\frac{\partial^2 F}{\partial x^2} = xy^2 g''(xy) + 2yg'(xy)$, $\frac{\partial^2 F}{\partial x \partial y} = x^2 y g''(xy) + 2xg'(xy)$, $\frac{\partial^2 F}{\partial y^2} = x^3 g''(xy)$

20. $\frac{\partial^2 F}{\partial x^2} = 4x^2(x^2+y^2)g''(x^2+y^2) + 2(5x^2+y^2)g'(x^2+y^2) + 2g(x^2+y^2)$

$\frac{\partial^2 F}{\partial x \partial y} = 4xy(x^2+y^2)g''(x^2+y^2) + 8xyg'(x^2+y^2)$

$\frac{\partial^2 F}{\partial y^2} = 4y^2(x^2+y^2)g''(x^2+y^2) + 2(5y^2+x^2)g'(x^2+y^2) + 2g(x^2+y^2)$

21. $\frac{\partial^2 F}{\partial x^2} = g'(g(x)+g(y))g''(x) + g''(g(x)+g(y))(g'(x))^2$

$\frac{\partial^2 F}{\partial x \partial y} = g'(x)g'(y)g''(g(x)+g(y))$

$\frac{\partial^2 F}{\partial y^2} = g'(g(x)+g(y))g''(y) + g''(g(x)+g(y))(g'(y))^2$

Respuestas a los ejercicios del capítulo 3,

Capítulo 3, Sección 1 (página 246)

1. **a.** $g(u, v) = (u, v)$, **b.** $F(u, v) = u$
2. **a.** $g(u, v) = (u^2 v, uv)$, **b.** $F(u, v) = \operatorname{sen}(u^2 v) + \operatorname{sen}(uv)$
3. **a.** $g(u, v) = (u + v, u)$, **b.** $F(u, v) = 3(u+v)^2 + 8u^3$
4. **a.** $g(u, v) = (\operatorname{sen} u, \cos u)$, **b.** $F(u, v) = \frac{1}{2}(\operatorname{sen} u + \cos u)$
5. **a.** $g(u, v) = (u - v + \pi/4, v - u)$, **b.** $F(u, v) = 1$
6. **a.** $F(x, y) = a^2 x + b(a+1)y$, **b.** $F(x, y) = a(b+1)x + b^2 y$, **c.** $F(x, y) = a^2 x + ab(x+y) + b^2 y$, **d.** $F(x, y) = (a^2 + b^2)x + 2by$;
7. **a.** $F(x, y) = x^4 y^3$, **b.** $F(x, y) = x^4 y$, **c.** $F(x, y) = x^6 y^3$, **d.** $F(x, y) = x^5 y^4$
8. **a.** $F(x, y) = \operatorname{sen}(\operatorname{sen} x + \operatorname{sen} y) + \operatorname{sen} y$, **b.** $F(x, y) = \operatorname{sen} x + \operatorname{sen}(\operatorname{sen} x + \operatorname{sen} y)$, **c.** $F(x, y) = 2\operatorname{sen}(\operatorname{sen} x + \operatorname{sen} y)$, **d.** $F(x, y) = 2\operatorname{sen}(\operatorname{sen} x + \operatorname{sen} y)$
9. **a.** $F(x, y) = \ln(1 + |1 + \ln|x||)$, **b.** $F(x, y) = \ln(1 + |x|)$, **c.** $F(x, y) = \ln(1 + |1 + \ln|x||)$, **d.** $F(x, y) = \ln(1 + |1 + \ln|x||)$
10. **a.** $F(x, y) = \arctan y$, **b.** $F(x, y) = \arctan(\arctan y)$, **c.** $F(x, y) = \arctan(\arctan y)$, **d.** $F(x, y) = \arctan(\arctan x)$
12. $F(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n)$. Dominio: $\{(x_1, x_2, \ldots, x_n) | x_1 x_2 \ldots x_n \neq 0\}$ (todo el espacio $\mathbb{R}^n$ excepto los "planos coordenados").
13. **a.** $(f \circ g)(u, v) = 3u + 2v$, **b.** $(h \circ g)(u, v) = 3u + 4v$, **c.** $f \circ g \circ h$ no se puede

14. **a.** $(f \circ g)(u, v) = u^3v^3$, **b.** $(h \circ g)(u, v) = uv$, **c.** $f \circ g \circ h$ no se puede
15. **a.** $(f \circ g)(u, v) = 2(u^2 + v^2)$, **b.** $h \circ g$ no se puede, **c.** $(f \circ g \circ h)(t) = 10t^2$
16. **a.** $f \circ g$ no se puede, **b.** $h \circ g$ no se puede, **c.** $f \circ g \circ h$ no se puede
17. **a.** $(f \circ g)(u) = 4u$, **b.** $h \circ g$ no se puede, **c.** $(f \circ g \circ h)(t) = 4t^2$
18. $(f \circ g)(u, v, w) = u^2v^2w^2$, **b.** $(h \circ g)(u, v, w) = (u^2v^2, u^3w^3, v^4w^4)$,
c. $(f \circ g \circ h)(r, s, t) = r^4s^6t^8$
19. $(f \circ g)(u, v, w) = \operatorname{sen} u$, **b.** $h \circ g$ no se puede, **c.** $(f \circ g \circ h)(t) = \operatorname{sen} t$
20. $(f \circ g)(u, v, w) = 1$, **b.** $(h \circ g)(u, v, w) = (w^2, uvw, u^4v^5)$, **c.** $(f \circ g \circ h)(r, s, t) = 1$
21. (Las respuestas en este ejercicio no son únicas) **a.** $f(x, y) = 3xy$, $g(u) = \operatorname{sen} u$
b. $f(x, y) = x^3 - \operatorname{sen} y$, $g(u) = \arctan^2 u$ **c.** $f(x, y) = 1 - x^2 - y^2$, $g(u) = \sqrt{u}$
d. $f(x, y) = x - y$, $g(u) = \ln u$
22. (Las respuestas en este ejercicio no son únicas) **a.** $f(u, v) = \operatorname{sen} u + \cos v$,
$g(x, y) = (x + y, x - y)$ **b.** $f(u, v) = \arctan^2 u + v$, $g(x, y) = (5x + y^2, 3x - 2y)$
c. $f(u, v) = u^3 + 4v^4$, $g(x, y) = (3x - y, 2x + y)$ **d.** $f(u, v) = \ln u + e^v$, $g(x, y) = (|xy|, x)$
23. (Las respuestas en este ejercicio no son únicas) **a.** $f(u, v, w) = u^2 + v^2 + w^2$,
$g(x, y, z) = (x, y, z)$ **b.** $f(u, v, w) = \operatorname{sen} u + \cos v + \operatorname{sen}^2 w$, $g(x, y, z) = (x + y, y - z, x + y + z)$
c. $f(u, v, w) = u^2 + 5v^4 - 10w$, $g(x, y, z) = (x + y + z, 3x - y, z)$
d. $f(u, v, w) = \ln(1 + u) + v$, $g(x, y, z) = (x^2 + y^2 + z^2, x + y + z, 1)$
24. $f(x, y, z) = xyz$, $g_1(t) = t^2$, $g_2(t) = t + 3$, $g_3(t) = \sqrt{t}$, $g_4(t) = \operatorname{sen} t$, $g_5(t) = t^4$, $g_6(t) = 8t$
$g_7(t) = t + 1$, $g_8(t) = \ln t$, $g_9(t) = t^2$, $g_{10}(t) = t + 5$, $g_{11}(t) = \sqrt{t}$, $g_{12}(t) = 5t$, $g_{13}(t) = t + 1$
26. $F(x, y) = (x, y)$
27. $F = f$
28. $F(x, y, z) = (0, 0, 0)$

Capítulo 3, Sección 2 (página 264)

1. $F(x, y) = f(u)$, $u = g(x, y) = ax + by$, $\frac{\partial F}{\partial x}(x, y) = af'(u)$, $\frac{\partial F}{\partial y}(x, y) = bf'(u)$
2. $F(x, y) = xf(u)$, $u = g(x, y) = xy$, $\frac{\partial F}{\partial x}(x, y) = xyf'(u) + f(u)$, $\frac{\partial F}{\partial y}(x, y) = x^2 f'(u)$
3. $F(x, y) = x^2 f(u) - yf(v)$, $u = g_1(x, y) = x \operatorname{sen} y$, $v = g_2(x, y) = 2xy$,
$\frac{\partial F}{\partial x}(x, y) = x^2 \operatorname{sen} y f'(u) + 2xf(u) - 2y^2 f'(v)$, $\frac{\partial F}{\partial y}(x, y) = x^3 \cos y f'(u) - 2xyf'(v) - f(v)$
4. $F(x, y) = f(u, v)$, $u = g_1(x, y) = 2$, $v = g_2(x, y) = xy$, $\frac{\partial F}{\partial x}(x, y) = y\frac{\partial f}{\partial v}(u, v)$,
$\frac{\partial F}{\partial y}(x, y) = x\frac{\partial f}{\partial v}(u, v)$
5. $F(x, y) = f(u, v)$, $u = g_1(x, y) = y$, $v = g_2(x, y) = x$, $\frac{\partial F}{\partial x}(x, y) = \frac{\partial f}{\partial v}(u, v)$,
$\frac{\partial F}{\partial y}(x, y) = \frac{\partial f}{\partial u}(u, v)$
6. $F(x, y) = f(x, y) + f(u, v)$, $u = g_1(x, y) = y^2$, $v = g_2(x, y) = x^2$,
$\frac{\partial F}{\partial x}(x, y) = \frac{\partial f}{\partial x}(x, y) + 2x\frac{\partial f}{\partial v}(u, v)$, $\frac{\partial F}{\partial y}(x, y) = \frac{\partial f}{\partial y}(x, y) + 2y\frac{\partial f}{\partial u}(u, v)$
7. $F(x, y) = f(u, v) + f(r, s)$, $u = s = g_1(x, y) = x - y$, $v = r = g_2(x, y) = x + y$
$\frac{\partial F}{\partial x}(x, y) = \frac{\partial f}{\partial u}(u, v) + \frac{\partial f}{\partial v}(u, v) + \frac{\partial f}{\partial r}(r, s) + \frac{\partial f}{\partial s}(r, s)$,
$\frac{\partial F}{\partial y}(x, y) = -\frac{\partial f}{\partial u}(u, v) + \frac{\partial f}{\partial v}(u, v) + \frac{\partial f}{\partial r}(r, s) - \frac{\partial f}{\partial s}(r, s)$,

8. $F(x, y, z) = f(u, v)$, $u = g_1(x, y, z) = x + y$, $v = g_2(x, y, z) = y - z$,
$\frac{\partial F}{\partial x}(x, y, z) = \frac{\partial f}{\partial u}(u, v)$, $\frac{\partial F}{\partial y}(x, y, z) = \frac{\partial f}{\partial u}(u, v) + \frac{\partial f}{\partial v}(u, v)$, $\frac{\partial F}{\partial z}(x, y, z) = -\frac{\partial f}{\partial v}(u, v)$

9. $F(x, y, z) = f(u_1, u_2, u_3) + f(v_1, v_2, v_3)$, $u_1 = v_3 = g_1(x, y, z) = x$,
$u_2 = v_2 = g_2(x, y, z) = xy$, $u_3 = v_1 = g_3(x, y, z) = xyz$,
$\frac{\partial F}{\partial x}(x, y, z) = \frac{\partial f}{\partial u_1}(u_1, u_2, u_3) + \frac{\partial f}{\partial v_3}(v_1, v_2, v_3) + y\left(\frac{\partial f}{\partial u_2}(u_1, u_2, u_3) + \frac{\partial f}{\partial v_2}(v_1, v_2, v_3)\right) +$
$yz\left(\frac{\partial f}{\partial u_3}(u_1, u_2, u_3) + \frac{\partial f}{\partial v_1}(v_1, v_2, v_3)\right)$
$\frac{\partial F}{\partial y}(x, y, z) = x\left(\frac{\partial f}{\partial u_2}(u_1, u_2, u_3) + \frac{\partial f}{\partial v_2}(v_1, v_2, v_3)\right) + xz\left(\frac{\partial f}{\partial u_3}(u_1, u_2, u_3) + \frac{\partial f}{\partial v_1}(v_1, v_2, v_3)\right)$
$\frac{\partial F}{\partial z}(x, y, z) = xy\left(\frac{\partial f}{\partial u_3}(u_1, u_2, u_3) + \frac{\partial f}{\partial v_1}(v_1, v_2, v_3)\right)$

10. $F(x, y, z) = f(u_1, u_2, u_3)$, $u_1 = g_1(x, y, z) = xy^2z^3$, $u_2 = g_2(x, y, z) = x^3y^2z$,
$u_3 = g_3(x, y, z) = x^2y^2z^2$,
$\frac{\partial F}{\partial x}(x, y, z) = y^2z^3\frac{\partial f}{\partial u_1}(u_1, u_2, u_3) + 3x^2y^2z\frac{\partial f}{\partial u_2}(u_1, u_2, u_3) + 2xy^2z^2\frac{\partial f}{\partial u_3}(u_1, u_2, u_3)$
$\frac{\partial F}{\partial y}(x, y, z) = 2xyz^3\frac{\partial f}{\partial u_1}(u_1, u_2, u_3) + 2x^3yz\frac{\partial f}{\partial u_2}(u_1, u_2, u_3) + 2x^2yz^2\frac{\partial f}{\partial u_3}(u_1, u_2, u_3)$
$\frac{\partial F}{\partial z}(x, y, z) = 3xy^2z^2\frac{\partial f}{\partial u_1}(u_1, u_2, u_3) + x^3y^2\frac{\partial f}{\partial u_2}(u_1, u_2, u_3) + 2x^2y^2z\frac{\partial f}{\partial u_3}(u_1, u_2, u_3)$

11. $F(x, y, z) = f(u_1, u_2, u_3, u_4)$, $u_1 = g_1(x, y, z) = x + 3y$, $u_2 = g_2(x, y, z) = 2y - 3z$,
$u_3 = g_3(x, y, z) = 2x + 7y - 6z$, $u_4 = g_4(x, y, z) = x - y - z$
$\frac{\partial F}{\partial x}(x, y, z) = \frac{\partial f}{\partial u_1}(u_1, u_2, u_3, u_4) + 2\frac{\partial f}{\partial u_3}(u_1, u_2, u_3, u_4) + \frac{\partial f}{\partial u_4}(u_1, u_2, u_3, u_4)$
$\frac{\partial F}{\partial y}(x, y, z) = 3\frac{\partial f}{\partial u_1}(u_1, u_2, u_3, u_4) + 2\frac{\partial f}{\partial u_2}(u_1, u_2, u_3, u_4) + 7\frac{\partial f}{\partial u_3}(u_1, u_2, u_3, u_4) -$
$\frac{\partial f}{\partial u_4}(u_1, u_2, u_3, u_4)$
$\frac{\partial F}{\partial z}(x, y, z) = -3\frac{\partial f}{\partial u_2}(u_1, u_2, u_3, u_4) - 6\frac{\partial f}{\partial u_3}(u_1, u_2, u_3, u_4) - \frac{\partial f}{\partial u_4}(u_1, u_2, u_3, u_4)$

12. $F(x, y, z) = f(u_1, u_2, u_3)$, $u_1 = g_1(x, y, z) = 1 - x$, $u_2 = g_2(x, y, z) = 2 - y$,
$u_3 = g_3(x, y, z) = 3 - z$,
$\frac{\partial F}{\partial x}(x, y, z) = -\frac{\partial f}{\partial u_1}(u_1, u_2, u_3)$, $\frac{\partial F}{\partial y}(x, y, z) = -\frac{\partial f}{\partial u_2}(u_1, u_2, u_3)$,
$\frac{\partial F}{\partial z}(x, y, z) = -\frac{\partial f}{\partial u_3}(u_1, u_2, u_3)$

13. $F(x, y, z, u) = f(v_1, v_2, v_3, v_4)$, $v_1 = g_1(x, y, z, u) = -x$, $v_2 = g_2(x, y, z, u) = -y$,
$v_3 = g_3(x, y, z, u) = -z$, $v_4 = g_4(x, y, z, u) = -u$
$\frac{\partial F}{\partial x}(x, y, z, u) = -\frac{\partial f}{\partial v_1}(v_1, v_2, v_3, v_4)$, $\frac{\partial F}{\partial y}(x, y, z, u) = -\frac{\partial f}{\partial v_2}(v_1, v_2, v_3, v_4)$
$\frac{\partial F}{\partial z}(x, y, z, u) = -\frac{\partial f}{\partial v_3}(v_1, v_2, v_3, v_4)$, $\frac{\partial F}{\partial u}(x, y, z, u) = -\frac{\partial f}{\partial v_4}(v_1, v_2, v_3, v_4)$

14. $F(x, y, z, u) = f(v_1, v_2, v_3, v_4)$, $v_1 = g_1(x, y, z, u) = \operatorname{sen} x, v_2 = g_2(x, y, z, u) = \cos y$,
$v_3 = g_3(x, y, z, u) = \tan z$, $v_4 = g_4(x, y, z, u) = \cot u$
$\frac{\partial F}{\partial x}(x, y, z, u) = \cos x\frac{\partial f}{\partial v_1}(v_1, v_2, v_3, v_4)$, $\frac{\partial F}{\partial y}(x, y, z, u) = -\operatorname{sen} y\frac{\partial f}{\partial v_2}(v_1, v_2, v_3, v_4)$

$\frac{\partial F}{\partial z}(x, y, z, u) = \sec^2 z \frac{\partial f}{\partial v_3}(v_1, v_2, v_3, v_4)$, $\frac{\partial F}{\partial u}(x, y, z, u) = -\csc^2 u \frac{\partial f}{\partial v_4}(v_1, v_2, v_3, v_4)$

15. $F(x_1, x_2, x_3, x_4) = f(v_1, v_2, v_3, v_4, v_5, v_6)$, $v_1 = x_1x_2$, $v_2 = x_1x_3$, $v_3 = x_1x_4$, $v_4 = x_2x_3$, $v_5 = x_2x_4$, $v_6 = x_3x_4$

$\frac{\partial F}{\partial x_1}(x_1, x_2, x_3, x_4) = x_2\frac{\partial f}{\partial v_1}(v_1, \ldots, v_6) + x_3\frac{\partial f}{\partial v_2}(v_1, \ldots, v_6) + x_4\frac{\partial f}{\partial v_3}(v_1, \ldots, v_6)$

$\frac{\partial F}{\partial x_2}(x_1, x_2, x_3, x_4) = x_1\frac{\partial f}{\partial v_1}(v_1, \ldots, v_6) + x_3\frac{\partial f}{\partial v_4}(v_1, \ldots, v_6) + x_4\frac{\partial f}{\partial v_5}(v_1, \ldots, v_6)$

$\frac{\partial F}{\partial x_3}(x_1, x_2, x_3, x_4) = x_1\frac{\partial f}{\partial v_2}(v_1, \ldots, v_6) + x_2\frac{\partial f}{\partial v_4}(v_1, \ldots, v_6) + x_4\frac{\partial f}{\partial v_6}(v_1, \ldots, v_6)$

$\frac{\partial F}{\partial x_4}(x_1, x_2, x_3, x_4) = x_1\frac{\partial f}{\partial v_3}(v_1, \ldots, v_6) + x_2\frac{\partial f}{\partial v_5}(v_1, \ldots, v_6) + x_3\frac{\partial f}{\partial v_6}(v_1, \ldots, v_6)$

16. $F(t) = f(u)$, $u = g(t) = 8t^2 + 13t - 1$, $F'(t) = (16t + 13)f'(u)$

17. $F(t) = f(u, v)$, $u = g_1(t) = 3t + 2$, $v = g_2(t) = 5t - 4$, $F'(t) = 3\frac{\partial f}{\partial u}(u, v) + 5\frac{\partial f}{\partial v}(u, v)$

18. $F(t) = (t^3 + 2)f(u, v)$, $u = g_1(t) = 1$, $v = g_2(t) = t$, $F'(t) = (t^3 + 2)\frac{\partial f}{\partial v}(u, v) + 3t^2 f(u, v)$

19. $F(t) = tf(u, v, w)$, $u = v = w = g(t) = t$,
$F'(t) = t\frac{\partial f}{\partial u}(u, v, w) + \frac{\partial f}{\partial v}(u, v, w) + \frac{\partial f}{\partial w}(u, v, w) + f(u, v, w)$

20. $F(t) = \operatorname{sen}^2 t^3 \cos f(u, v)$, $u = g_1(t) = \operatorname{sen} t$, $v = g_2(t) = \cos t$,
$F'(t) = -\operatorname{sen}^2 t^3 \operatorname{sen} f(u, v) \cos t \frac{\partial f}{\partial u}(u, v) - \operatorname{sen} t \frac{\partial f}{\partial v}(u, v) + 3t^2 \operatorname{sen}(2t^3) \cos f(u, v)$

23. $F'(0) = 4$

24. $\operatorname{grad}(f \circ g)(0, 0) = (9, -32)$

25. $\operatorname{grad} F(1, 1, \ldots, 1) = \left(\sum_{i=1}^{n} a_i, \sum_{i=2}^{n} a_i, \ldots, a_{n-1} + a_n, a_n\right)$

26. $\operatorname{grad} F(1, 1, \ldots, 1) = \left(\sum_{i=1}^{n} i, \sum_{i=2}^{n} i, \ldots, 2n - 1, n\right)$

28. $13/\sqrt{2}$

29. $b(5\alpha_1 + 7\alpha_2)$

30. En la dirección del vector (22, 17)

31. $z = 3x + 6y$

32. **a.** $x - 2y + z = 0$

37. $F(x, y) = f(u)$, $u = g(x, y) = ax + by$

$\frac{\partial^2 F}{\partial x^2}(x, y) = a^2 f''(u)$, $\frac{\partial^2 F}{\partial x \partial y}(x, y) = abf''(u)$, $\frac{\partial^2 F}{\partial y^2}(x, y) = b^2 f''(u)$

38. $F(x, y) = f(u, v)$, $u = g_1(x, y) = ax + by$, $v = g_2(x, y) = cx + dy$,

$\frac{\partial^2 F}{\partial x^2}(x, y) = a^2\frac{\partial^2 f}{\partial u^2}(u, v) + 2ac\frac{\partial^2 f}{\partial u \partial v}(u, v) + c^2\frac{\partial^2 f}{\partial v^2}(u, v)$

$\frac{\partial^2 F}{\partial x \partial y}(x, y) = ab\frac{\partial^2 f}{\partial u^2}(u, v) + (ad + bc)\frac{\partial^2 f}{\partial u \partial v}(u, v) + cd\frac{\partial^2 f}{\partial v^2}(u, v)$

$\frac{\partial^2 F}{\partial y^2}(x, y) = b^2\frac{\partial^2 f}{\partial u^2}(u, v) + 2bd\frac{\partial^2 f}{\partial u \partial v}(u, v) + d^2\frac{\partial^2 f}{\partial v^2}(u, v)$

39. $\frac{\partial^2 F}{\partial x^2}(x, y) = y\frac{\partial^2 f}{\partial x^2}(x, x) + 2\frac{\partial^2 f}{\partial x \partial y}(x, x) + \frac{\partial^2 f}{\partial y^2}(x, x)$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = \frac{\partial f}{\partial x}(x, x) + \frac{\partial f}{\partial y}(x, x) + \frac{\partial f}{\partial x}(y, y) + \frac{\partial f}{\partial y}(y, y)$$

$$\frac{\partial^2 F}{\partial y^2}(x, y) = x\frac{\partial^2 f}{\partial x^2}(y, y) + 2\frac{\partial^2 f}{\partial x \partial y}(y, y) + \frac{\partial^2 f}{\partial y^2}(y, y)$$

40. $F(x, y) = (x^2 + y^2) f(u, v)$, $u = g_1(x, y) = x$, $v = g_2(x, y) = 2y$,

$$\frac{\partial^2 F}{\partial x^2}(x, y) = (x^2 + y^2)\frac{\partial^2 f}{\partial u^2}(u, v) + 4x\frac{\partial f}{\partial u}(u, v) + 2f(u, v)$$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = 2(x^2 + y^2)\frac{\partial^2 f}{\partial u \partial v}(u, v) + 2y\frac{\partial f}{\partial u}(u, v) + 4x\frac{\partial f}{\partial v}(u, v)$$

$$\frac{\partial^2 F}{\partial y^2}(x, y) = 4(x^2 + y^2)\frac{\partial^2 f}{\partial v^2}(u, v) + 8y\frac{\partial f}{\partial v}(u, v) + 2f(u, v)$$

41. $F(x, y) = f(u, v, w)$, $u = x$, $v = xy$, $w = x + y$

$$\frac{\partial^2 F}{\partial x^2}(x, y) = \frac{\partial^2 f}{\partial u^2}(u, v, w) + 2y\frac{\partial^2 f}{\partial u \partial v}(u, v, w) + 2\frac{\partial^2 f}{\partial u \partial w}(u, v, w) + y^2\frac{\partial^2 f}{\partial v^2}(u, v, w)$$

$$+ 2y\frac{\partial^2 f}{\partial v \partial w}(u, v, w) + \frac{\partial^2 f}{\partial w^2}(u, v, w)$$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = x\frac{\partial^2 f}{\partial u \partial v}(u, v, w) + \frac{\partial^2 f}{\partial u \partial w}(u, v, w) + \frac{\partial f}{\partial v}(u, v, w) + y\left(x\frac{\partial^2 f}{\partial v^2}(u, v, w) + \right.$$

$$\left.\frac{\partial^2 f}{\partial v \partial w}(u, v, w)\right) + x\frac{\partial^2 f}{\partial v \partial w}(u, v, w) + \frac{\partial^2 f}{\partial w^2}(u, v, w)$$

$$\frac{\partial^2 F}{\partial y^2}(x, y) = x^2\frac{\partial^2 f}{\partial v^2}(u, v, w) + 2x\frac{\partial^2 f}{\partial v \partial w}(u, v, w) + \frac{\partial^2 f}{\partial w^2}(u, v, w)$$

42. $\frac{\partial^2 F}{\partial x^2}(x, y) = g(f(x, y))\frac{\partial^2 f}{\partial x^2}(x, y) + \left(\frac{\partial f}{\partial x}(x, y)\right)^2 g'(f(x, y)) - g'(x)$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = g(f(x, y))\frac{\partial^2 f}{\partial x \partial y}(x, y) + \frac{\partial f}{\partial x}(x, y)\frac{\partial f}{\partial y}(x, y)g'(f(x, y))$$

$$\frac{\partial^2 F}{\partial y^2}(x, y) = g'(f(x, y))\frac{\partial f}{\partial y}(x, y)$$

43. Sean $u = x + y$, $v = x - y$

$$\frac{\partial^2 F}{\partial x^2}(x, y) =$$

$$g(f(u, v))\frac{\partial^2 f}{\partial u^2}(u, v) + 2\frac{\partial^2 f}{\partial u \partial v}(u, v) + \frac{\partial^2 f}{\partial v^2}(u, v) + g'(f(u, v))\frac{\partial f}{\partial u}(u, v) + \frac{\partial f}{\partial v}(u, v)^2 - y^2 g'(xy)$$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = g(f(u, v))\frac{\partial^2 f}{\partial u^2}(u, v) - \frac{\partial^2 f}{\partial v^2}(u, v) + \frac{\partial f}{\partial u}(u, v)^2 - \frac{\partial f}{\partial v}(u, v)^2 g'(f(u, v)) - xyg'(xy) - g(xy)$$

$$\frac{\partial^2 F}{\partial y^2}(x, y) =$$

$$g(f(u, v))\frac{\partial^2 f}{\partial u^2}(u, v) - 2\frac{\partial^2 f}{\partial u \partial v}(u, v) + \frac{\partial^2 f}{\partial v^2}(u, v) + g'(f(u, v))\frac{\partial f}{\partial u}(u, v) - \frac{\partial f}{\partial v}(u, v)^2 - x^2 g'(xy)$$

44. Sean $u = \operatorname{sen} x$, $v = \cos y$, $w = xy$

$$\frac{\partial^2 F}{\partial x^2}(x, y) = g'(x) - g(yf(u, v, w))\left(y\cos^2 x\frac{\partial^2 f}{\partial u^2}(u, v, w) + 2y^2\cos x\frac{\partial^2 f}{\partial u \partial w}(u, v, w)\right.$$

$$\left.+ y^3\frac{\partial^2 f}{\partial w^2}(u, v, w)\right) + \left(y\cos x\frac{\partial f}{\partial u}(u, v, w) + y^2\frac{\partial f}{\partial w}(u, v, w)\right)^2 g'(yf(u, v, w))$$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y) = -g(yf(u, v, w))\left[\cos x \frac{\partial f}{\partial u}(u, v, w) + y\cos x\left(-\operatorname{sen} y \frac{\partial^2 f}{\partial u \partial v}(u, v, w)\right.\right.$$
$$\left. + x\frac{\partial^2 f}{\partial u \partial w}(u, v, w)\right) + y^2\left(-\operatorname{sen} y \frac{\partial^2 f}{\partial v \partial w}(u, v, w) + x\frac{\partial^2 f}{\partial w^2}(u, v, w)\right) + 2y\frac{\partial f}{\partial w}(u, v, w)\Bigg] +$$
$$\left(y\cos x \frac{\partial f}{\partial u}(u, v, w) + y^2\frac{\partial f}{\partial w}(u, v, w)\right)g'(yf(u, v, w))\Bigg[f(u, v, w)$$
$$\left. + y\left(-\operatorname{sen} y\frac{\partial f}{\partial v}(u, v, w) + x\frac{\partial f}{\partial w}(u, v, w)\right)\right]$$
$$\frac{\partial^2 F}{\partial y^2}(x, y) = -g(yf(u, v, w))\left(-2\operatorname{sen} y\frac{\partial f}{\partial u}(u, v, w) + 2x\frac{\partial f}{\partial w}(u, v, w) - y\cos y\frac{\partial f}{\partial v}(u, v, w) +\right.$$
$$\left. y\operatorname{sen}^2 y\frac{\partial^2 f}{\partial v^2}(u, v, w) - 2xy\operatorname{sen} y\frac{\partial^2 f}{\partial v \partial w}(u, v, w) + x^2 y\frac{\partial^2 f}{\partial w^2}(u, v, w)\right) - g'(yf(u, v, w))\left(f(u, v, w) -\right.$$
$$\left. y\operatorname{sen} y\frac{\partial f}{\partial v}(u, v, w) + xy\frac{\partial f}{\partial w}(u, v, w)^2\right)$$

45. $F(x, y, z) = f(u),\ u = g(x, y, z) = ax + by + cz,$

$$\frac{\partial^2 F}{\partial x^2}(x, y, z) = a^2 f''(u),\ \frac{\partial^2 F}{\partial x \partial y}(x, y, z) = abf''(u),\ \frac{\partial^2 F}{\partial y^2}(x, y) = b^2 f''(u),$$
$$\frac{\partial^2 F}{\partial z^2}(x, y, z) = c^2 f''(u),\ \frac{\partial^2 F}{\partial x \partial z}(x, y, z) = acf''(u),\ \frac{\partial^2 F}{\partial y \partial z}(x, y, z) = bcf''(u)$$

46. $F(x, y, z) = xf(u, v),\ u = g_1(x, y, z) = xy,\ v = g_2(x, y, z) = xz,$

$$\frac{\partial^2 F}{\partial x^2}(x, y, z) = 2\left(y\frac{\partial f}{\partial u}(u, v) + z\frac{\partial f}{\partial v}(u, v)\right) + x\left(y^2\frac{\partial^2 f}{\partial u^2}(u, v) + 2yz\frac{\partial^2 f}{\partial u \partial v}(u, v) + z^2\frac{\partial^2 f}{\partial v^2}(u, v)\right)$$
$$\frac{\partial^2 F}{\partial x \partial y}(x, y, z) = 2x\frac{\partial f}{\partial u}(u, v) + x^2 y\frac{\partial^2 f}{\partial u^2}(u, v) + x^2 z\frac{\partial^2 f}{\partial u \partial v}(u, v)$$
$$\frac{\partial^2 F}{\partial x \partial z}(x, y, z) = 2x\frac{\partial f}{\partial v}(u, v) + x^2 y\frac{\partial^2 f}{\partial u \partial v}(u, v) + x^2 z\frac{\partial^2 f}{\partial v^2}(u, v)$$
$$\frac{\partial^2 F}{\partial y^2}(x, y, z) = x^3\frac{\partial^2 f}{\partial u^2}(u, v),\ \frac{\partial^2 F}{\partial y \partial z}(x, y, z) = x^3\frac{\partial^2 f}{\partial u \partial v}(u, v)$$
$$\frac{\partial^2 F}{\partial z^2}(x, y, z) = x^3\frac{\partial^2 f}{\partial v^2}(u, v)$$

47. $F(x, y, z) = f(u, v, w),\ u = g_1(x, y, z) = xyz,\ v = g_2(x, y, z) = xy, w = g_3(x, y, z) = x,$

$$\frac{\partial^2 F}{\partial x^2}(x, y, z) = y^2 z^2\frac{\partial^2 f}{\partial u^2}(u, v, w) + 2y^2 z\frac{\partial^2 f}{\partial u \partial v}(u, v, w) + 2yz\frac{\partial^2 f}{\partial u \partial w}(u, v, w) + y^2\frac{\partial^2 f}{\partial v^2}(u, v, w) +$$
$$2y\frac{\partial^2 f}{\partial v \partial w}(u, v, w) + \frac{\partial^2 f}{\partial w^2}(u, v, w)$$
$$\frac{\partial^2 F}{\partial x \partial y}(x, y, z) = xyz^2\frac{\partial^2 f}{\partial u^2}(u, v, w) + 2xyz\frac{\partial^2 f}{\partial u \partial v}(u, v, w) + z\frac{\partial f}{\partial u}(u, v, w) + \frac{\partial f}{\partial v}(u, v, w)$$
$$+ xy\frac{\partial^2 f}{\partial v^2}(u, v, w) + xz\frac{\partial^2 f}{\partial u \partial w}(u, v, w) + x\frac{\partial^2 f}{\partial v \partial w}(u, v, w)$$
$$\frac{\partial^2 F}{\partial x \partial z}(x, y, z) = xy^2 z\frac{\partial^2 f}{\partial u^2}(u, v, w) + xy^2\frac{\partial^2 f}{\partial u \partial v}(u, v, w) + y\frac{\partial f}{\partial u}(u, v, w) + xy\frac{\partial^2 f}{\partial u \partial w}(u, v, w)$$
$$\frac{\partial^2 F}{\partial y^2}(x, y, z) = x^2 z^2\frac{\partial^2 f}{\partial u^2}(u, v, w) + 2x^2 z\frac{\partial^2 f}{\partial u \partial v}(u, v, w) + x^2\frac{\partial^2 f}{\partial v^2}(u, v, w)$$

$$\frac{\partial^2 F}{\partial y \partial z}(x, y, z) = x^2 yz \frac{\partial^2 f}{\partial u^2}(u, v, w) + x \frac{\partial f}{\partial u}(u, v, w) + x^2 y \frac{\partial^2 f}{\partial u \partial v}(u, v, w),$$

$$\frac{\partial^2 F}{\partial z^2}(x, y, z) = x^2 y^2 \frac{\partial^2 f}{\partial u^2}(u, v, w)$$

48. $$\frac{\partial^2 F}{\partial x^2}(x, y, z) = \frac{\partial f}{\partial x}(f(x, y, z), y, z) \frac{\partial^2 f}{\partial x^2}(x, y, z) + \left(\frac{\partial f}{\partial x}(x, y, z)\right)^2 \frac{\partial^2 f}{\partial x^2}(f(x, y, z), y, z)$$

$$\frac{\partial^2 F}{\partial x \partial y}(x, y, z) = \frac{\partial f}{\partial x}(f(x, y, z), y, z) \frac{\partial^2 f}{\partial x \partial y}(x, y, z) + \frac{\partial f}{\partial x}(x, y, z)\left(\frac{\partial f}{\partial y}(x, y, z) \frac{\partial^2 f}{\partial x^2}(f(x, y, z), y, z) + \frac{\partial^2 f}{\partial x \partial y}(f(x, y, z), y, z)\right)$$

$$\frac{\partial^2 F}{\partial x \partial z}(x, y, z) = \frac{\partial f}{\partial x}(f(x, y, z), y, z) \frac{\partial^2 f}{\partial x \partial z}(x, y, z) + \frac{\partial f}{\partial x}(x, y, z)\left(\frac{\partial f}{\partial z}(x, y, z) \frac{\partial^2 f}{\partial x^2}(f(x, y, z), y, z) + \frac{\partial^2 f}{\partial x \partial z}(f(x, y, z), y, z)\right)$$

$$\frac{\partial^2 F}{\partial y^2}(x, y, z) = \frac{\partial f}{\partial x}(f(x, y, z), y, z) \frac{\partial^2 f}{\partial y^2}(x, y, z) + \frac{\partial^2 f}{\partial y^2}(f(x, y, z), y, z) + \left(\frac{\partial f}{\partial y}(x, y, z)\right)^2 \frac{\partial^2 f}{\partial x^2}(f(x, y, z), y, z) + 2 \frac{\partial f}{\partial y}(x, y, z) \frac{\partial^2 f}{\partial x \partial y}(f(x, y, z), y, z)$$

$$\frac{\partial^2 F}{\partial y \partial z}(x, y, z) = \frac{\partial f}{\partial x}(f(x, y, z), y, z) \frac{\partial^2 f}{\partial y \partial z}(x, y, z) + \frac{\partial f}{\partial y}(x, y, z)\left(\frac{\partial f}{\partial z}(x, y, z) \frac{\partial^2 f}{\partial x^2}(f(x, y, z), y, z) + \frac{\partial^2 f}{\partial x \partial z}(f(x, y, z), y, z)\right) + \frac{\partial f}{\partial z}(x, y, z) \frac{\partial^2 f}{\partial x \partial y}(f(x, y, z), y, z) + \frac{\partial^2 f}{\partial y \partial z}(f(x, y, z), y, z)$$

$$\frac{\partial^2 F}{\partial z^2}(x, y, z) = \frac{\partial f}{\partial x}(f(x, y, z), y, z) \frac{\partial^2 f}{\partial z^2}(x, y, z) + \left(\frac{\partial f}{\partial z}(x, y, z)\right)^2 \frac{\partial^2 f}{\partial x^2}(f(x, y, z), y, z) + 2 \frac{\partial f}{\partial z}(x, y, z) \frac{\partial^2 f}{\partial x \partial z}(f(x, y, z), y, z) + \frac{\partial^2 f}{\partial z^2}(f(x, y, z), y, z)$$

49. $F(x, y) = f(u, v)$, $u = g_1(x, y) = ax + by$, $v = g_2(x, y) = cx + dy$,

$$\frac{\partial^3 F}{\partial x^3}(x, y) = \left(a \frac{\partial}{\partial u} + c \frac{\partial}{\partial v}\right)^3 f(u, v),$$

$$\frac{\partial^3 F}{\partial x^2 \partial y}(x, y) = a^2 b \frac{\partial^3 f}{\partial u^3}(u, v) + (a^2 d + 2abc) \frac{\partial^3 f}{\partial u^2 \partial v}(u, v) + (2acd + bc^2) \frac{\partial^3 f}{\partial v^2 \partial u}(u, v) + c^2 d \frac{\partial^3 f}{\partial v^3}(u, v)$$

$$\frac{\partial^3 F}{\partial y^2 \partial x}(x, y) = ab^2 \frac{\partial^3 f}{\partial u^3}(u, v) + (b^2 c + 2\text{abd}) \frac{\partial^3 f}{\partial u^2 \partial v}(u, v) + (2bcd + ad^2) \frac{\partial^3 f}{\partial v^2 \partial u}(u, v) + cd^2 \frac{\partial^3 f}{\partial v^3}(u, v)$$

$$\frac{\partial^3 F}{\partial y^3}(x, y) = \left(b \frac{\partial}{\partial u} + d \frac{\partial}{\partial v}\right)^3 f(u, v)$$

Capítulo 3, Sección 3 (página 276)

1. $[0\ 0]$
2. $[a\ b\ c]$
3. $\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$
4. $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 0 & -1 \\ 4 & 4 & 4 \end{bmatrix}$

7. $[1\ 2\ 3\ 4]$

8. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

10. $\begin{bmatrix} 2 & 0 \\ 2\ln 2 & 1 \\ 2e^2 & e^2 \\ e^2 & e^2 \\ 2e^2 & e \end{bmatrix}$

12. $(14, -15)$

13. $(1/2, -1/2)$

14. $y = f(0, 0, 0) + 3x$

15. **b.** $10(x - x_0) + 31(y - y_0) + 30(z - z_0) = 0$, en donde $p = (x_0, y_0, z_0)$, **c.** $-140/\sqrt{5}$

16. $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

18. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

19. $\begin{bmatrix} 27 & 10 \\ 5 & 2 \end{bmatrix}$

20. $\begin{bmatrix} 145 & 54 \\ 27 & 10 \end{bmatrix}$

21. $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix}$

22. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

23. $\begin{bmatrix} 7 & 0 \\ 5 & 2 \end{bmatrix}$

24. $[5\ 1]$

25. $[0\ 0]$

26. $[3\ -1]$

27. $[0\ 20]$

28. $[0\ 0]$

29. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 3 & -1 \end{bmatrix}$

30. $\begin{bmatrix} 3 & -1 \\ 2 & 2 \\ 5 & 1 \end{bmatrix}$

31. $\begin{bmatrix} 1 & -2 \\ 6 & 8 \\ 0 & 0 \end{bmatrix}$

32. $\begin{bmatrix} 11 & -1 \\ 39 & 11 \\ 2 & 10 \end{bmatrix}$

33. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

34. $g(x, y) = (f_2(x, y), f_1(x, y))$
35. $g(x, y) = (2f_1(x, y), 3f_2(x, y))$
36. $g(x, y) = (x, f_1(x, y))$
37. $g(x, y) = (4x + 5y, 8f_2(x, y))$
38. $g(x, y) = (f_1(x, y) + 3f_2(x, y), f_2(x, y) - f_1(x, y))$
39. $g(x, y) = (4f_1(x, y) + 5x - 12y, f_1(x, y) + f_2(x, y) + 3x - 10y)$
40. $g(x, y) = (0.5f_1^2(x, y), 0.5f_2^2(x, y))$
41. $g = f \circ f \circ \ldots \circ f$ (k veces)

Capítulo 3, Sección 4 (página 294)

1. $y'(x) = -4/5$
2. $y'(x) = -2y/(2x + 1)$
3. $y'(x) = -(9x^2 + 2x + 8x^3)/24xy^2$
4. $y'(x) = 5e^{x-y} - e^{-y}$
5. $y'(x) = 4(4x^3 + 4x^2 + 3y^3x - 17x + 3y^3 + 1)/81y^2$
6. $y'(\mathbf{p}) = 8/5$
7. $y'(\mathbf{p}) = -1$
8. $y'(\mathbf{p}) = 1$
9. $y'(\mathbf{p}) = -1$
10. $y'(\mathbf{p}) = -1$
11. $\frac{\partial f}{\partial x}(\mathbf{p}) = -1,\ \frac{\partial f}{\partial y}(\mathbf{p}) = 3$
12. $\frac{\partial f}{\partial x}(\mathbf{p}) = 0,\ \frac{\partial f}{\partial y}(\mathbf{p}) = 0$
13. $\frac{\partial f}{\partial x}(\mathbf{p}) = -1,\ \frac{\partial f}{\partial y}(\mathbf{p}) = 8$
14. $\frac{\partial f}{\partial x}(\mathbf{p}) = -1/3,\ \frac{\partial f}{\partial y}(\mathbf{p}) = -1/6$
15. $\frac{\partial f}{\partial x}(\mathbf{p}) = 3/e,\ \frac{\partial f}{\partial y}(\mathbf{p}) = -3/e$
19. $f'(x) = \dfrac{\frac{\partial F}{\partial y}(y, x) - \frac{\partial F}{\partial x}(x, y)}{\frac{\partial F}{\partial y}(x, y) - \frac{\partial F}{\partial x}(y, x)}$

20. $$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, \varphi(x), y) + \frac{\partial F}{\partial y}(x, \varphi(x), y)\varphi'(x)}{\frac{\partial F}{\partial z}(x, \varphi(x), y)}$$

21. $$f'(x) = -\frac{\frac{\partial F}{\partial x}(\varphi(x), \psi(x), y^3)\varphi'(x) + \frac{\partial F}{\partial y}(\varphi(x), \psi(x), y^3)\psi'(x)}{3y^2\frac{\partial F}{\partial z}(\varphi(x), \psi(x), y^3)}$$

22. $$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, y)\frac{\partial G}{\partial x}(F(x, y), F(y, x)) + \frac{\partial F}{\partial y}(y, x)\frac{\partial G}{\partial y}(F(x, y), F(y, x))}{\frac{\partial F}{\partial y}(x, y)\frac{\partial G}{\partial x}(F(x, y), F(y, x)) + \frac{\partial F}{\partial x}(y, x)\frac{\partial G}{\partial y}(F(x, y), F(y, x))}$$

23. **a.** Sí es posible verla como una función del tipo $u = u(x, y, z)$. Sus derivadas parciales son $\frac{\partial u}{\partial x}(1, 1, 1) = \frac{\partial u}{\partial y}(1, 1, 1) = \frac{\partial u}{\partial z}(1, 1, 1) = -1$.
b. Sí es posible verla como una función del tipo $z = z(x, y, u)$. Sus derivadas parciales son $\frac{\partial z}{\partial x}(1, 1, 1) = \frac{\partial z}{\partial y}(1, 1, 1) = \frac{\partial z}{\partial u}(1, 1, 1) = -1$.
c. Sí es posible verla como una función del tipo $y = y(x, z, u)$. Sus derivadas parciales son $\frac{\partial y}{\partial x}(1, 1, 1) = \frac{\partial y}{\partial z}(1, 1, 1) = \frac{\partial y}{\partial u}(1, 1, 1) = -1$.
d. Sí es posible verla como una función del tipo $x = x(y, z, u)$. Sus derivadas parciales son $\frac{\partial x}{\partial y}(1, 1, 1) = \frac{\partial x}{\partial z}(1, 1, 1) = \frac{\partial x}{\partial u}(1, 1, 1) = -1$.

24. **a.** Sí es posible verla como una función del tipo $u = u(x, y, z)$. Sus derivadas parciales son $\frac{\partial u}{\partial x}(0, 0, 1) = 0$, $\frac{\partial u}{\partial y}(0, 0, 1) = 5/8$, $\frac{\partial u}{\partial z}(0, 0, 1) = 1$.
b. Sí es posible verla como una función del tipo $z = z(x, y, u)$. Sus derivadas parciales son $\frac{\partial z}{\partial x}(0, 0, 1) = 0$, $\frac{\partial z}{\partial y}(0, 0, 1) = -5/8$, $\frac{\partial z}{\partial u}(0, 0, 1) = 1$.
c. Sí es posible verla como una función del tipo $y = y(x, z, u)$. Sus derivadas parciales son $\frac{\partial y}{\partial x}(0, 0, 1) = 0$, $\frac{\partial y}{\partial z}(0, 0, 1) = -8/5$, $\frac{\partial y}{\partial u}(0, 0, 1) = 8/5$. **d.** No es posible verla como una función del tipo $x = x(y, z, u)$.

25. **a.** Sí es posible verla como una función del tipo $u = u(x, y, z)$. Sus derivadas parciales son $\frac{\partial u}{\partial x}(0, 0, 0) = 0$, $\frac{\partial u}{\partial y}(0, 0, 0) = 8$, $\frac{\partial u}{\partial z}(0, 0, 0) = 3/8$.
b. Sí es posible verla como una función del tipo $z = z(x, y, u)$. Sus derivadas parciales son $\frac{\partial z}{\partial x}(0, 0, 0) = 0$, $\frac{\partial z}{\partial y}(0, 0, 0) = 0$, $\frac{\partial z}{\partial u}(0, 0, 0) = 8/3$.
c. No es posible verla como una función del tipo $y = y(x, z, u)$.
d. No es posible verla como una función del tipo $x = x(y, z, u)$.

26. No es posible verla como ninguna de las funciones $u = u(x, y, z)$, $z = z(x, y, u)$, $y = y(x, z, u)$, $x = x(y, z, u)$.

27. $2/\sqrt{5}$

28. $\sqrt{3}$

29. En la dirección del vector $(-1, -1)$

30. En cualquier punto de la superficie distinto del $(16, 8, 0)$

31. En cualquier punto de la superficie distinto del $(1, 3, -4)$

32. $\dfrac{\partial^2 f}{\partial x^2}(x, y) = \dfrac{2y(32x^2y^2z + 3(8yz^2 - 1)^2)}{9(1 - 8yz^2)^3}$

$\dfrac{\partial^2 f}{\partial x \partial y}(x, y) = \dfrac{2x(16x^2y^2z + 128y^2z^4 - 24yz^3 + 3)}{9(1 - 8yz^2)^3}$

$\dfrac{\partial^2 f}{\partial y^2}(x, y) = \dfrac{16z(x^2y + z(3 - 16yz^2))(x^2 + 8z^3)}{9(1 - 8yz^2)^3}$

33. $\dfrac{\partial^2 f}{\partial x^2}(x, y) = \dfrac{y^2(\operatorname{sen} z \cos^2(xy) + \operatorname{sen}(xy)(1 + \cos z)^2)}{(1 + \cos z)^3}$

$\dfrac{\partial^2 f}{\partial x \partial y}(x, y) = \dfrac{xy \operatorname{sen} z \cos^2(xy) - (\cos z + 1)^2(xy \operatorname{sen}(xy) - \cos(xy))}{(1 + \cos z)^3}$

$\dfrac{\partial^2 f}{\partial y^2}(x, y) = \dfrac{x^2(\operatorname{sen} z \cos^2(xy) + \operatorname{sen}(xy)(1 + \cos z)^2)}{(1 + \cos z)^3}$

34. $\dfrac{\partial^2 f}{\partial x^2}(1, 1) = \dfrac{\partial^2 f}{\partial y^2}(1, 1) = -2,\ \dfrac{\partial^2 f}{\partial x \partial y}(1, 1) = -3/2$

37. $\dfrac{\partial f}{\partial x}(x, y) = -\dfrac{D_2F(y, x, z)}{D_3F(y, x, z)},\ \dfrac{\partial f}{\partial y}(x, y) = -\dfrac{D_1F(y, x, z)}{D_3F(y, x, z)}$ en donde D_iF es la derivada parcial de F respecto de su i-ésima variable.

38. $\dfrac{\partial f}{\partial x}(x, y) = -\dfrac{D_3F(z, y, x)}{D_1F(z, y, x)},\ \dfrac{\partial f}{\partial y}(x, y) = -\dfrac{D_2F(z, y, x)}{D_1F(z, y, x)}$

39. $\dfrac{\partial f}{\partial x}(x, y) = -\dfrac{D_1F(x, y, z) + D_1F(x, z, y) + D_2F(z, x, y)}{D_3F(x, y, z) + D_2F(x, z, y) + D_1F(z, x, y)}$

$\dfrac{\partial f}{\partial y}(x, y) = -\dfrac{D_2F(x, y, z) + D_3F(x, z, y) + D_3F(z, x, y)}{D_3F(x, y, z) + D_2F(x, z, y) + D_1F(z, x, y)}$

40. $\dfrac{\partial f}{\partial x}(x, y) = -\dfrac{a_1D_1(u, v, w) + a_2D_2(u, v, w) + a_3D_3(u, v, w)}{c_1D_1(u, v, w) + c_2D_2(u, v, w) + c_3D_3(u, v, w)}$

$\dfrac{\partial f}{\partial y}(x, y) = -\dfrac{b_1D_1(u, v, w) + b_2D_2(u, v, w) + b_3D_3(u, v, w)}{c_1D_1(u, v, w) + c_2D_2(u, v, w) + c_3D_3(u, v, w)}$ en donde $u = a_1x + b_1y + c_1z$, $v = a_2x + b_2y + c_2z$, $w = a_3x + b_3y + c_3z$

41. $\dfrac{\partial f}{\partial x}(x, y) =$

$$-\frac{F(y, y, z) + y(D_1F(x, x, z) + D_2F(x, x, z)) + z(D_1F(x, x, x) + D_2F(x, x, x) + D_3F(x, x, x))}{xD_3F(y, y, z) + yD_3F(x, x, z) + F(x, x, x)}$$

$\dfrac{\partial f}{\partial y}(x, y) = -\dfrac{xD_1F(y, y, z) + xD_2F(y, y, z) + F(x, x, z)}{xD_3F(y, y, z) + yD_3F(x, x, z) + F(x, x, x)}$

42. $\dfrac{\partial f}{\partial x}(x, y) =$

$$\frac{-\operatorname{sen} y \cos z D_1F(u, v, w) + z \cos x \cos y D_3F(u, v, w)}{-x \operatorname{sen} y \operatorname{sen} z D_1F(u, v, w) + y \cos z \cos y D_2F(u, v, w) + \operatorname{sen} x \cos y D_3F(u, v, w)}$$

$\dfrac{\partial f}{\partial y}(x, y) = -\dfrac{+x \cos y \cos z D_1F(u, v, w) + (\operatorname{sen} z \cos y - y \operatorname{sen} z \operatorname{sen} y)D_2F(u, v, w)}{-x \operatorname{sen} y \operatorname{sen} z D_1F(u, v, w) + y \cos z \cos y D_2F(u, v, w) + \operatorname{sen} x \cos y D_3F(u, v, w)}$ en donde $u = x \operatorname{sen} y \cos z$, $v = y \operatorname{sen} z \cos y$, $w = z \operatorname{sen} x \cos y$

43. $\dfrac{\partial f}{\partial x}(x, y) = -\dfrac{y^{-1}D_1F(u, v, w) - zx^{-2}D_3F(u, v, w)}{x^{-1}D_3F(u, v, w) - yz^{-2}D_2F(u, v, w)}$

$\dfrac{\partial f}{\partial y}(x, y) = -\dfrac{z^{-1}D_2F(u, v, w) - xy^{-2}D_1F(u, v, w)}{x^{-1}D_3F(u, v, w) - yz^{-2}D_2F(u, v, w)}$ en donde $u = xy^{-1}$, $v = yz^{-1}$, $w = zx^{-1}$

44. $\dfrac{\partial^2 f}{\partial x^2}(x, y) = -\dfrac{\left(\dfrac{\partial F}{\partial v}\right)^2 \dfrac{\partial^2 F}{\partial u^2} - 2\dfrac{\partial F}{\partial u}\dfrac{\partial F}{\partial v}\dfrac{\partial^2 F}{\partial u \partial v} + \left(\dfrac{\partial F}{\partial u}\right)^2 \dfrac{\partial^2 F}{\partial v^2}}{\left(\dfrac{\partial F}{\partial u}\right)^3}$ en donde $u = x + z$, $v = x$,

todas las derivadas de F evaluadas en (u, v).

$\dfrac{\partial^2 f}{\partial x \partial y}(x, y) = \dfrac{\partial^2 f}{\partial y^2}(x, y) = 0$ (observe que la función f no depende de y).

45. $\dfrac{\partial f}{\partial x}(x, y) = -\dfrac{3h(3x + y) + zg(xz)}{g(y + z) - xg(xz) + 2zh(z^2)}$

$\dfrac{\partial f}{\partial y}(x, y) = -\dfrac{g(y + z) - h(3x + y)}{g(y + z) - xg(xz) + 2zh(z^2)}$

Capítulo 3, Sección 5 (página 305)

2. $\dfrac{\partial u}{\partial x}(1, 1) = 2 - e$, $\dfrac{\partial v}{\partial x}(1, 1) = 1 - e^{-1}$, $\dfrac{\partial u}{\partial y}(1, 1) = e$, $\dfrac{\partial v}{\partial y}(1, 1) = 0$.

3. Plano tangente a $u = u(x, y)$ en $(1, 1, 1)$: $7x - 5y - 4z + 2 = 0$

Plano tangente a $v = v(x, y)$ en $(1, 1, 1)$: $5x - 3y - 4z + 2 = 0$.

8. $\dfrac{\partial u}{\partial x}(0, 1) = 1$, $\dfrac{\partial v}{\partial x}(0, 1) = 0$, $\dfrac{\partial u}{\partial y}(0, 1) = 0$, $\dfrac{\partial v}{\partial y}(0, 1) = 1$

9. Si $\Delta = D_1 f(u, v) D_2 g(u, v) - D_2 f(u, v) D_1 g(u, v) \neq 0$, se tiene

$\dfrac{\partial u}{\partial x}(x, y) = \dfrac{1}{\Delta}(D_1 g(x, y) D_2 g(u, v) - D_2 f(u, v) D_1 f(x, y))$

10. Si $\Delta = D_1 g(u, v) D_1 g(v, u) - D_2 g(u, v) D_2 g(v, u) \neq 0$, se tiene

$\dfrac{\partial v}{\partial x}(x, y) = \dfrac{1}{\Delta}(D_1 f(x, y) D_2 g(v, u) - D_1 g(u, v) D_2 f(y, x))$

11. Si $\Delta = D_1 f(u, x) D_2 g(y, v) - D_2 f(x, u) D_1 g(v, y) \neq 0$, se tiene

$\dfrac{\partial u}{\partial y}(x, y) = \dfrac{1}{\Delta}(D_2 g(v, y) D_2 g(y, v) - D_1 g(y, v) D_1 g(v, y))$

12. Si $\Delta = D_1 f(u, v) D_2 g(u, v) - D_2 f(u, v) D_1 g(u, v) \neq 0$, se tiene

$\dfrac{\partial u}{\partial y}(x, y) = \dfrac{1}{\Delta}(D_2 f(x, y) D_2 g(u, v) - D_2 f(u, v) D_2 g(x, y))$

13. Si $\Delta = g(u, u) g(v, v) - uv(D_1 g(v, v) + D_2 g(v, v))(D_1 g(u, u) + D_2 g(u, u)) \neq 0$

$\dfrac{\partial u}{\partial x}(x, y) = \dfrac{1}{\Delta}(yg(u, u)(D_1 f(x, x) + D_2 f(x, x)) - uf(y, y)(D_1 g(v, v) + D_2 g(v, v)))$

14. Si $\Delta = (xf(y, v) - D_1 g(u, v))(yg(x, u) - D_2 f(u, v)) - xuD_2 f(y, v) -$
$D_2 g(u, v))(yvD_2 g(x, u) - D_1 f(u, v)) \neq 0$, se tiene

$\dfrac{\partial v}{\partial x}(x, y) = \dfrac{1}{\Delta}(uf(y, v)(yvD_2 g(x, u) - D_1 f(u, v)) - yvD_1 g(x, u)(xf(y, v) - D_1 g(u, v)))$

15. Si $\Delta = (xD_1 f(xu, yv) - D_1 g(u, v))(xD_2 f(yu, xv) - D_2 g(u, v))$
$- (yD_2 f(xu, yv) - D_2 g(u, v))(yD_1 f(yu, xv) - D_1 g(u, v)) \neq 0$, se tiene

$\dfrac{\partial u}{\partial x}(x, y) = \dfrac{1}{\Delta}(vD_2 f(yu, xv)(yD_2 f(xu, yv) - D_2 g(u, v))$
$- uD_1 f(xu, yv)(xD_2 f(yu, xv) - D_2 g(u, v)))$

16. $\dfrac{\partial^2 u}{\partial x \partial y}(0, 0) = 0$, $\dfrac{\partial^2 v}{\partial x \partial y}(0, 0) = 0$

17. $\dfrac{\partial u}{\partial x}(1, 1) = 0$, $\dfrac{\partial v}{\partial x}(1, 1) = -1/2$, $\dfrac{\partial u}{\partial y}(1, 1) = 0$, $\dfrac{\partial v}{\partial y}(1, 1) = 1/2$

18. $\dfrac{\partial u}{\partial x}(1, 1, 1) = 7$, $\dfrac{\partial v}{\partial x}(1, 1, 1) = -1$, $\dfrac{\partial u}{\partial y}(1, 1, 1) = 3$, $\dfrac{\partial v}{\partial y}(1, 1, 1) = 0$

$\frac{\partial u}{\partial z}(1,1,1) = 7,\ \frac{\partial v}{\partial z}(1,1,1) = -1$

19. $\frac{\partial u}{\partial x} = 3,\ \frac{\partial v}{\partial y} = 2,\ \frac{\partial w}{\partial z} = 0$

20. $\frac{\partial w}{\partial x} = 2,\ \frac{\partial v}{\partial y} = 2/3,\ \frac{\partial u}{\partial z} = 1/3,\ \frac{\partial x}{\partial u} = 1/3,\ \frac{\partial y}{\partial v} = 2/3,\ \frac{\partial z}{\partial w} = 0$

Capítulo 3, Sección 6 (página 317)

1. $JF^{-1}(F(\mathbf{p})) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$

2. $JF^{-1}(F(\mathbf{p})) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

3. $JF^{-1}(F(\mathbf{p})) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

4. $JF^{-1}(F(\mathbf{p})) = \begin{bmatrix} (2\operatorname{sen} 1)^{-1} & -(2\operatorname{sen} 1)^{-1} \\ (2\cos 1)^{-1} & (2\cos 1)^{-1} \end{bmatrix}$

5. $JF^{-1}(F(\mathbf{p})) = \begin{bmatrix} 4/3 & -2/3 \\ -2/3 & 4/3 \end{bmatrix}$

6. $F^{-1}(x, y) = (x/a, y/b)$, $\det JF(x, y) = ab$, $\det JF^{-1}(F(x, y)) = (ab)^{-1}$

10. $JF^{-1}(3, 2) = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$, $x = x(u, v) = \frac{1}{2}(u + \sqrt{u^2 - 4v})$,

$y = y(u, v) = \frac{1}{2}(u - \sqrt{u^2 - 4v})$

11. $\operatorname{grad} x(0, 0) = (2/7, -1/7)$, $\operatorname{grad} y(0, 0) = (1/7, 3/7)$

12. $JF^{-1}(0, 0) = \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}$

13. $F^{-1}(x, y, z) = (x - y, y - z, z)$

17. $F^{-1} = F$

Capítulo 3, Sección 7 (página 331)

1. Con $A = 3$: $(\pm 0.9233145, 0.1214456, 0.3643368)$;
con $A = 1$: $(\pm 0.8570399, 0.3643368, 0.3643368)$;
con $A = 0.391226858$: $(\pm 0.0000574, 0.9312673, 0.3643368)$.

2. $(1.2686457013, -0.7313542987, 0.2686457013)$.

3. $(0.1994413116, 0.1655761846)$.

4. $\mathbf{p}_1 = (\pm 0.8931356229, 0.4497874598)$; $\mathbf{p}_2 = (0.4497874598, \pm 0.8931356229)$.

5. $\mathbf{p}_1 = (1.16690991, 0.9094460719)$; $\mathbf{p}_2 = (-0.1319809628, 0.7355746769)$;
$\mathbf{p}_3 = (-1.4602034134, 1.5182494323)$.

6. $\mathbf{p}_1 = (1.4314539066, 1, 1.0490602866)$; $\mathbf{p}_2 = (-0.6039125638, -1, -0.6352896152)$;
$\mathbf{p}_3 = (-0.9314539066, 1, -0.1323936199)$; $\mathbf{p}_4 = (1.1039125638, -1, 0.2186229486)$.

7. $(0.539178639, 0.1247533122, 0.6639319512)$.

8. $\mathbf{p}_1 = (2.3470626915, 2.2597489385, 0.7541801719)$;
$\mathbf{p}_2 = (0.9656150103, -4.3566402185, -0.950833089)$;
$\mathbf{p}_3 = (-0.9287753817, -4.4779445557, 0.9617685776)$.

9. $(2.8280781933, -2.754658272, -1.1972866953)$.

10. $\mathbf{p}_1 = (0.1739078762, 1.7574393849, 0.6991035281)$;
$\mathbf{p}_2 = (0.0461893325, 1.9158101415, 0.9116254496)$

Respuestas a los ejercicios del capítulo 4,

Capítulo 4, Sección 1 (página 340)

28. $(4, 3/2)$
29. $(-1/2, 0)$
30. $(-1, 2)$
31. No hay puntos críticos
32. $(1, 1)$, $(2, 1)$
33. $(5/2, 0)$
34. No hay puntos críticos
35. $(0, \pi/2 + k\pi)$, $k \in \mathbb{Z}$
36. No hay puntos críticos
37. $(0, 0, 0)$
38. $(0, 0, 0)$
39. $(-1/2, -1/2, -1/2)$
40. $(1, 2, t)$, $(1, t, 3)$, $(t, 2, 3)$, $t \in \mathbb{R}$
41. $(2, -1, t)$, $t \in \mathbb{R}$

Capítulo 4, Sección 2 (página 352)

1. $f((x_0, y_0) + (x, y)) = 5$
2. $f((x_0, y_0) + (x, y)) = 5x_0 + 5x$
3. $f((x_0, y_0) + (x, y)) = 5x_0 + 8y_0 + 4 + 5x + 8y$
4. $f((x_0, y_0) + (x, y)) = 10y_0^2 + 20y_0y + 10y^2$
5. $f((x_0, y_0) + (x, y)) = 3x_0y_0 + 3y_0x + 3x_0y + 3xy$
6. $f((x_0, y_0)+(x, y)) = 2x_0^2+7x_0y_0+5y_0^2-2+(4x_0+7y_0)x+(7x_0+10y_0)y+2x^2+7xy+5y^2$
7. $f((x_0, y_0)+(x, y)) = x_0^3+y_0^3+3x_0^2x+3y_0^2y+3x_0x^2+3y_0y^2+r(x, y)$ en donde $r(x, y) = x^3+y^3$
8. $f((x_0, y_0) + (x, y)) = 3x_0 - 2y_0 + 15z_0 - 23 + 3x - 2y + 15z$
9. $f((x_0, y_0) + (x, y)) = 2x_0^2 - 5y_0^2 + 3z_0^2 + x_0y_0 - 6x_0z_0 + 2y_0z_0 + 1 + (4x_0 + y_0 - 6z_0)x + (-10y_0 + x_0 + 2z_0)y + (6z_0 - 6x_0 + 2y_0)z + 2x^2 - 5y^2 + 3z^2 + xy - 6xz + 2yz$
14. **a.** $3x^2+4y^2-8xy+5 = 8-10(x-1)+8(y-2)+3(x-1)^2-8(x-1)(y-2)+4(y-2)^2$
b. $5x^2 - 10y^2 + 14 = -31 - 30(x+3) - 60(y-3) + 5(x+3)^2 - 10(y-3)^2$
c. $x^2 + y^2 = 169 + 10(x-5) + 24(y-12) + (x-5)^2 + (y-12)^2$
15. **a.** $x^2 + y^2 + z^2 = 14 + 2(x-1) + 4(y-2) + 6(y-3) + (x-1)^2 + (y-2)^2 + (z-3)^2$
b. $2x^2+3y^2-5z^2+3xy+3 = 1-3x-6(y+1)-10(z-1)+2x^2+3x(y+1)+3(y+1)^2-5(z-1)^2$
c. $x^2 - 2y^2 + 4z^2 + xy + xz - yz + 1 = -1 + (x-1) + 5(y+1) + 2z + (x-1)^2 - 2(y+1)^2 - 4z^2 + (x-1)(y+1) + (x-1)z - (y+1)z$
20. $x^3 + y^3 = r(x, y)$
21. $\dfrac{1}{1+x+y} = 1 - x - y + x^2 + y^2 + 2xy + r(x, y)$
22. $\dfrac{1}{1+x^2+y^2} = 1 - x^2 - y^2 + r(x, y)$
23. $\ln(1-x) + \ln(1-y) = -x - y - x^2 - y^2 + r(x, y)$
24. $e^x \operatorname{sen} y = y + xy + r(x, y)$
25. $e^x \cos y = 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2 + r(x, y)$
26. $\arctan(x+y) = x + y + r(x, y)$
27. $z \approx 1 - \frac{3}{2}(x-1) - (y-1) - \frac{39}{8}(x-1)^2 - \frac{39}{4}(x-1)(y-1) - 2(y-1)^2$

$f(1.1, 0.9) \approx 0.97875$, $f(0.912, 1.087) \approx 1.066756$

Capítulo 4, Sección 3 (página 361)

1. Mínimo local
2. Punto de ensilladura
3. No se puede concluir nada
4. Punto de ensilladura
5. Máximo local
6. Mínimo local
7. Punto de ensilladura
8. Punto de ensilladura
9. Punto de ensilladura
10. Mínimo local
11. **a.** $f(x, y) = \frac{1}{2}(x^2 + y^2)$, $\mathbf{p} = (0, 0)$,
b. $f(x, y) = x^2 + y^2 + 4xy$, $\mathbf{p} = (0, 0)$;
12. **a.** $f(x, y, z) = \frac{1}{2}(x^2 + y^2 + z^2)$, $\mathbf{p} = (0, 0, 0)$,
b. $f(x, y, z) = \frac{1}{2}x^2 - y^2 + 3xy + \frac{1}{2}z^2$, $\mathbf{p} = (0, 0, 0)$
15. Mínimo local.
16. Mínimo local.
17. Mínimo local.
18. Máximo local.
19. Máximo local.
20. Punto de ensilladura.
21. Máximo local.
22. Máximo local.
23. Punto de ensilladura.
24. No se puede concluir nada.
25. Mínimo local.
26. Mínimo local en $\mathbf{p} = (-11/7, -5/7)$, $f(\mathbf{p}) = -15/7$.
27. Punto de ensilladura en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 0$.
28. Punto de ensilladura en $\mathbf{p} = (-9/32, -5/32)$, $f(\mathbf{p}) = -175/64$.
29. Punto de ensilladura en $\mathbf{p} = (-1, -1)$, $f(\mathbf{p}) = -4$.
30. Puntos de ensilladura en $\mathbf{p}_1 = (1, 1)$ y $\mathbf{p}_2 = (-1, -1)$, $f(\mathbf{p}_1) = f(\mathbf{p}_2) = -1$; Mínimo local en $\mathbf{p}_3 = (1, -1)$, $f(\mathbf{p}_3) = -5$; Máximo local en $\mathbf{p}_4 = (-1, 1)$, $f(\mathbf{p}_4) = 3$.
31. Puntos de ensilladura en $\mathbf{p}_1 = (2^{-1/2}, 5/2)$ y $\mathbf{p}_2 = (-2^{-1/2}, 5/2)$, $f(\mathbf{p}_1) = f(\mathbf{p}_2) = 4$; Máximo local en $\mathbf{p}_3 = (0, 5/2)$, $f(\mathbf{p}_3) = 17/4$.
32. Punto de ensilladura en $\mathbf{p} = (0, 1/e)$, $f(\mathbf{p}) = 0$.
33. Punto de ensilladura en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 0$.
34. Máximo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = \ln 3$.
35. Mínimo local en $\mathbf{p} = (-1, -1, -2)$, $f(\mathbf{p}) = -53$.
36. Máximo local en $\mathbf{p} = (1, 1, 1)$, $f(\mathbf{p}) = 10$.
37. Punto de ensilladura en $\mathbf{p} = (-(5/2)^{1/3}, -(3/2)^{1/3}, (9/4)^{1/3})$, $f(\mathbf{p}) = 27(18)^{1/3}/8 - 15(20)^{1/3}/4 - 9(12)^{1/3}/2 \approx -11.63649$.
38. Punto de ensilladura en $\mathbf{p}_1 = (1, 1, 1)$, $f(\mathbf{p}_1) = 10/3$; Máximo local en $\mathbf{p}_2 = (-1, 1, 1)$, $f(\mathbf{p}_2) = 14/3$.

39. No hay extremos locales.
40. Punto de ensilladura en $\mathbf{p}_1 = (0, 1, 3/2)$, $f(\mathbf{p}_1) = 17/2$; Máximo local en $\mathbf{p}_2 = (-2, 1, 3/2)$, $f(\mathbf{p}_2) = 25/2$.
41. No hay extremos locales.
42. Punto de ensilladura en $\mathbf{p}_1 = (-1/3, -1/4, 1)$, $f(\mathbf{p}_1) = -481/72$; Mínimo local en $\mathbf{p}_2 = (1/15, -1/4, -1)$, $f(\mathbf{p}_2) = -12889/1800$.
43. Puntos de ensilladura en $\mathbf{p}_1 = (1, -2, -1)$ y $\mathbf{p}_2 = (-1, 0, -3)$, $f(\mathbf{p}_1) = -8$, $f(\mathbf{p}_2) = -4$.
44. Punto de ensilladura en $\mathbf{p}_1 = (-1/7, 10/49, -26/49)$, $f(\mathbf{p}_1) = -433/343$; Mínimo local en $\mathbf{p}_2 = (1/3, 0, -2/3)$, $f(\mathbf{p}_2) = -37/27$.
45. Puntos de ensilladura en $\mathbf{p}_1 = (0, 0, 0)$ y $\mathbf{p}_2 = (-4/3, -16/3, 8)$, $f(\mathbf{p}_1) = 0$, $f(\mathbf{p}_2) = 64/27$.
46. Mínimo local en $\mathbf{p} = (-1/4, -1/4, -1/4)$, $f(\mathbf{p}) = 13/8$.
47. Máximo local en $\mathbf{p} = (6/7, 10/7, 1/2)$, $f(\mathbf{p}) = 85/28$.
48. Mínimo local en $\mathbf{p} = (3/37, -17/37, -52/37)$, $f(\mathbf{p}) = -187/37$.
49. Mínimo local en $\mathbf{p} = (0, 0, 0)$, $f(\mathbf{p}) = 0$.
50. Mínimo local en $\mathbf{p} = (0, -1/2, 0)$, $f(\mathbf{p}) = 3/2$.
51. Máximo local en $\mathbf{p} = (0, -1/2, 0)$, $f(\mathbf{p}) = -3/2$.
52. Mínimo local en $\mathbf{p} = (0, -1/5, -2/5)$, $f(\mathbf{p}) = 1/5$.
53. Máximo local en $\mathbf{p} = (0, 1/2, 0)$, $f(\mathbf{p}) = 1/4$.
54. Máximo local en $\mathbf{p} = (1/3, 1/3, 2/3)$, $f(\mathbf{p}) = 8/3$.
55. Punto de ensilladura en $\mathbf{p} = (1/62, -21/62, -21/62)$, $f(\mathbf{p}) = -83/124$.
56. Punto de ensilladura en $\mathbf{p} = (-11/96, -53/96, -95/96)$, $f(\mathbf{p}) = 365/64$.
57. **a.** La función tiene un punto de ensilladura en (x_0, x_0), en donde vale 0.
b. La función tiene un punto de ensilladura en $(-x_0, x_0)$, en donde vale 0.
c. La función tiene un punto de ensilladura en $(x_0, -x_0)$, en donde vale 0.
d. La función tiene un punto de ensilladura en $(-x_0, -x_0)$, en donde vale 0.
58. **a.** La función tiene un mínimo local en (1, 1), que vale $f(1, 1) = 0$
b. La función tiene un máximo local en (1,1), que vale $f(1, 1) = 0$
59. **a.** La función tiene un mínimo local en (1, 1), que vale $f(1, 1) = 0$
b. La función tiene un máximo local en (1, 1), que vale $f(1, 1) = 0$
60. **a.** No se puede concluír nada sobre la naturaleza del punto crítico.
b. La función tiene un punto de ensilladura en (1, 1), en donde vale $f(1, 1) = \int_1^2 g(t)\,dt$.
61. La función tiene un punto de ensilladura en el origen, en donde es igual a cero.
64. Mínimo local en el punto $\mathbf{p}_1 = (-0.5586983783, 0.7120680454)$, en donde f vale -2.9164986198. Puntos de ensilladura en $\mathbf{p}_2 = (0.4801231603, 1.0705634765)$, en donde f vale -1.34662560689, y en $\mathbf{p}_3 = (-1.3428853954, -0.467782692)$ en donde f vale -1.50796628444.
65. Máximo local en el punto $\mathbf{p}_1 = (5.9967428306, 4.4780411506)$ en donde f vale -0.6050360514. Punto de ensilladura en $\mathbf{p}_2 = (1.9037042778, 1.60950402)$ en donde f vale -0.8133257086.
66. Mínimo local en $\mathbf{p} = (0.1994413116, 0.1655761846)$ en donde f vale 0.657974873.
67. Máximos locales en $\mathbf{p}_1 = (1.2608355088, 1.135866801)$, $\mathbf{p}_2 = (0.1739459925, -2.1288488911)$, $\mathbf{p}_3 = (-1.1632403062, 2.7896045616)$, en donde f vale $f(\mathbf{p}_1) = 2.020997077$, $f(\mathbf{p}_2) = 1.857469209875$, $f(\mathbf{p}_3) = 3.5862955307983$.
Mínimos locales en $\mathbf{q}_1 = (2.5324568988, -0.9988585036)$, $\mathbf{q}_2 = (-0.8427981441, 0.0605444921)$, $\mathbf{q}_3 = (-2.5575868237, -2.5430563921)$, en donde f vale $f(\mathbf{q}_1) = -3.3407267097$, $f(\mathbf{q}_2) = -0.56435863$, $f(\mathbf{q}_3) = -4.9110970201$.

Puntos de ensilladura en $r_1 = (0.4348247405, -0.375200907)$,
$r_2 = (-1.26878597, -0.5280065436)$, $r_3 = (0.3355947528, 1.6186551184)$,
$r_4 = (2.5576923509, 2.1716122186)$, $r_5 = (0.6223373491, -2.5762182087)$, en donde f vale
$f(r_1) = 0.2774452357$, $f(r_2) = -0.4217113823$, $f(r_3) = 1.5963410717$, $f(r_4) = 0.1981248059$,
$f(r_5) = 1.7681559234$.

68. Máximo local en $\mathbf{p} = (6.6525296152, 8.0164195703)$ en donde f vale 15.1896384702.
Mínimos locales en $\mathbf{q}_1 = (9.5162348028, 10.9951360308)$, $\mathbf{q}_2 = (3.3609598633, 4.7053427186)$,
$\mathbf{q}_3 = (9.6439219351, 4.7099382543)$, $\mathbf{q}_4 = (3.2330565938, 10.9942841151)$, en donde f vale
$f(\mathbf{q}_1) = 9.470378082$, $f(\mathbf{q}_2) = 3.2494444691$, $f(\mathbf{q}_3) = 9.5325755249$, $f(\mathbf{q}_4) = 3.1871945514$.
Puntos de ensilladura en $r_1 = (2.2054618601, 1.7262761895)$,
$r_2 = (5.6312029328, 4.5288431674)$, $r_3 = (9.2960600817, 7.8548678895)$,
$r_4 = (3.0129061499, 7.8567147757)$, $r_5 = (6.0103229102, 10.8138528073)$, en donde f vale
$f(r_1) = 2.7684246123$, $f(r_2) = 0.881913883075$, $f(r_3) = 9.3607702$, $f(r_4) = 3.0775925026$,
$f(r_5) = 5.8346869025$.

69. Mínimo local en (0.427756813 , 0.3258051706 , -0.0201042087) en donde f vale -0.0012075324.

Capítulo 4, Sección 4 (página 376)

1. Mínimo local en $\mathbf{p} = (36/23, -29/23)$, $f(\mathbf{p}) = -798/26$.

2. Punto de ensilladura en $\mathbf{p} = (-1, -1)$, $f(\mathbf{p}) = 0$.

3. Punto de ensilladura en $\mathbf{p} = (-1, -1)$, $f(\mathbf{p}) = -2$.

4. Puntos de ensilladura en $\mathbf{p}_1 = (1, -1)$ y $\mathbf{p}_2 = (-1, 1)$, $f(\mathbf{p}_1) = f(\mathbf{p}_2) = -2$; Mínimo local en $\mathbf{p}_3 = (1, 1)$, $f(\mathbf{p}_3) = -6$; Máximo local en $\mathbf{p}_4 = (-1, -1)$, $f(\mathbf{p}_4) = 2$.

5. Puntos de ensilladura en $\mathbf{p}_1 = (0, 1/\sqrt{3})$, $\mathbf{p}_2 = (1, -1/\sqrt{3})$, $\mathbf{p}_3 = (-1, -1/\sqrt{3})$, $f(\mathbf{p}_1) = -2/\sqrt{3} - 1$, $f(\mathbf{p}_2) = f(\mathbf{p}_3) = 2/\sqrt{3} - 2$; Mínimos locales en $\mathbf{p}_4 = (1, 1/\sqrt{3})$ y $\mathbf{p}_5 = (-1, 1/\sqrt{3})$, $f(\mathbf{p}_4) = f(\mathbf{p}_5) = -2/\sqrt{3} - 2$; Máximo local en $\mathbf{p}_6 = (0, -1/\sqrt{3})$, $f(\mathbf{p}_6) = 2/\sqrt{3} - 1$.

6. Punto de ensilladura en $\mathbf{p}_1 = (0, 0)$, $f(\mathbf{p}_1) = 0$; Máximo local en $\mathbf{p}_2 = (1/3, 1/3)$, $f(\mathbf{p}_2) = 1/27$.

7. Mínimo local en $\mathbf{p}_1 = (-1/\sqrt{2}, 0)$, $f(\mathbf{p}_1) = -(2e)^{-1/2}$; Máximo local en $\mathbf{p}_2 = (1/\sqrt{2}, 0)$, $f(\mathbf{p}_2) = (2e)^{-1/2}$.

8. Mínimo local en $\mathbf{p}_1 = (0, -1/\sqrt{2})$, $f(\mathbf{p}_1) = -(2e)^{-1/2}$; Máximo local en $\mathbf{p}_2 = (0, 1/\sqrt{2})$, $f(\mathbf{p}_2) = (2e)^{-1/2}$.

9. Punto de ensilladura en $\mathbf{p} = (0, 1)$, $f(\mathbf{p}) = 0$.

10. Punto de ensilladura en $\mathbf{p} = (0, e)$, $f(\mathbf{p}) = 0$.

11. Punto de ensilladura en $\mathbf{p} = (1, 4/3)$, $f(\mathbf{p}) = -2$.

12. Puntos de ensilladura en $\mathbf{p}_{1,2} = (1, \pm\sqrt{2})$, $f(\mathbf{p}_{1,2}) = -1$.

13. Mínimo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 0$.

14. Punto de ensilladura en $\mathbf{p} = (\tan 1, 2\sec^2 1)$. $f(\mathbf{p}) = 1 - 2\tan 1$.

15. Mínimo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 2$.

16. Puntos de ensilladura en puntos del tipo $((2k_1 - 1)\pi, 2k_2\pi)$ y $(2k_1\pi, (2k_2 - 1)\pi)$, $k_1, k_2 \in \mathbb{Z}$, en donde la función vale 0; Máximos locales en puntos del tipo $(2k_1\pi, 2k_2\pi)$, $k_1, k_2 \in \mathbb{Z}$, en donde la función vale 2; Mínimos locales en puntos del tipo $((2k_1 - 1)\pi),(2k_2 - 1)\pi)$, $k_1, k_2 \in \mathbb{Z}$, en donde la función vale -2.

17. Máximos locales en puntos del tipo $\left((\frac{\pi}{2} + 2k_1\pi), 2k_2\pi\right)$ y $\left((\frac{3\pi}{2} + 2k_1\pi), (2k_2 - 1)\pi\right)$, $k_1, k_2 \in \mathbb{Z}$ en donde la función vale 1.

Mínimos locales en puntos del tipo $\left((\frac{\pi}{2}+2k_1\pi), (2k_2-1)\pi\right)$ y $\left((\frac{3\pi}{2}+2k_1\pi), 2k_2\pi\right)$, k_1, $k_2 \in \mathbb{Z}$ en donde la función vale -1.

Puntos de ensilladura en puntos del tipo $\left(k_1\pi, (\frac{\pi}{2}+k_2\pi)\right)$, $k_1, k_2 \in \mathbb{Z}$ en donde la función vale 0.

18. Punto de ensilladura en $\mathbf{p} = (-1, 1)$, $f(\mathbf{p}) = -1$.
19. Punto de ensilladura en $\mathbf{p} = (1, 1)$, $f(\mathbf{p}) = 3/2$.
20. Mínimo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = -1$.
21. Punto de ensilladura en $\mathbf{p} = (0, -\ln 1)$, $f(\mathbf{p}) = -1 + \ln 1$.
22. Mínimo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 0$.
23. Máximo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 2$.
24. Máximo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 2$.
25. Mínimo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = \ln 2$.
26. Máximo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = \ln 2$.
27. Mínimo local en $\mathbf{p} = (0, 0)$, $f(\mathbf{p}) = 1/\ln 3$.
28. Puntos de ensilladura en $\mathbf{p}_1 = (-1, -17/2, 5)$ y $\mathbf{p}_2 = (1, 15/2, -7)$.
29. Puntos de ensilladura en $\mathbf{p}_1 = (-1/3, -7/9, 1)$ y $\mathbf{p}_2 = (1/3, 7/9, -1)$.
30. Máximo local en $\mathbf{p}_1 = (5/2, -3, -7/2)$; Mínimo local en $\mathbf{p}_2 = (-3, 3, 4)$.
31. Máximo local en $\mathbf{p}_1 = (-1/3, -2/3, 1/3)$; Mínimo local en $\mathbf{p}_2 = (1, 2, -5)$.
32. Máximo local en $\mathbf{p}_1 = (-7/6, -1/6, 53/6)$; Mínimo local en $\mathbf{p}_2 = -p_1$.
33. Máximo local en $\mathbf{p}_1 = (-1/2, -1/2, 0)$; Mínimo local en $\mathbf{p}_2 = (-3/8, -5/16, -3/4)$.
34. Mínimo local en $\mathbf{p}_1 = (0, 9, 5)$; Máximo local en $\mathbf{p}_2 = (0, 15, 9)$.
35. Puntos de ensilladura en $\mathbf{p}_1 = (2, -5/2, 1/2)$ y $\mathbf{p}_2 = (2, -5, -2)$.
36. Puntos de ensilladura en $\mathbf{p}_1 = (1, -9, -16)$ y $\mathbf{p}_2 = (-1, 11, 18)$.
37. **a.** La función tiene un mínimo local en (x_0, x_1), en donde f vale $a+b$
b1. La función tiene un mínimo local en (x_0, x_1), en donde f vale ab
b2. La función tiene un punto de ensilladura en (x_0, x_1), en donde f vale ab
b3. La función tiene un punto de ensilladura en (x_0, x_1), en donde f vale ab
b4. La función tiene un máximo local en (x_0, x_1), en donde f vale ab.
c1. La función tiene un mínimo local en (x_0, x_1), en donde f vale a^2+b^2
c2. La función tiene un punto de ensilladura en (x_0, x_1), en donde f es a^2+b^2
c3. La función tiene un punto de ensilladura en (x_0, x_1), en donde f es a^2+b^2
c4. La función tiene un máximo local en (x_0, x_1), en donde f vale a^2+b^2
40. $y = 0.9610796x + 0.0981234$
41. $y = 3.136008x - 1.3112791$
42. $y = 2.4983099x + 2.8816073$
44. **a.** $z = x + y + 1$; **b.** $z = 0.7070738x - 1.2599703y + 1.5196239$
c. $z = 0.4909552x - 0.5024791y + 1.5317031$;
d. $z = -0.9249261x + 3.0492042y - 1.1221165$;
e. $z = -1.2847011x + 0.4687284y - 1.4559002$
45. **a.** $y = 2.0090909x^2 - 1.0054545x + 2.9727273$ **b.** $y = -2.0035714x^2 + 8.995x - 0.9928571$
c. $y = 0.9756158x^2 - 2.9392857x + 4.0722906$
46. **a.** $y = 3x_1 + 2x_2 - x_3 - 4$
b. $y = 2.738478x_1 - 4.286662x_2 + 12.002445x_3 - 7.342877x_4 - 1.700978$

Capítulo 4, Sección 5 (página 402)

1. Máx. en $\mathbf{p} = (2, -3/2)$, $f(\mathbf{p}) = 11/2$. Mín. en $\mathbf{q} = (-2, 3/2)$ $f(\mathbf{q}) = -11/2$.

2. Máx. en $\mathbf{p} = (1, 6)$, $f(\mathbf{p}) = 19$. Mín. en $\mathbf{q} = (-1, -6)$, $f(\mathbf{q}) = -19$.
3. Máx. en $\mathbf{p} = (3, -28/3)$, $f(\mathbf{p}) = 223/3$. Mín. en $\mathbf{q} = (-3, 28/3)$, $f(\mathbf{q}) = -223/3$
4. Máx. en $\mathbf{p} = (-2, -15)$, $f(\mathbf{p}) = 79$. Mín. en $\mathbf{q} = (2, 15)$, $f(\mathbf{q}) = -79$.
5. Máx. en $\mathbf{p} = (1, 2)$, $f(\mathbf{p}) = 5$. Mín. en $\mathbf{q} = (-1, -2)$, $f(\mathbf{q}) = -5$.
6. Máximo $= S^2/4$, para $x = y = S/2$.
7. Máximo $= S^3/27$, para $x = y = z = S/3$.
8. Máximo en los puntos $\mathbf{p}_{1,2,3,4} = (\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$, $f(\mathbf{p}_{1,2,3,4}) = 1/4$. Mínimos en los puntos $\mathbf{p}_{5,6} = (\pm 1, 0)$, $\mathbf{p}_{7,8} = (0, \pm 1)$, $f(\mathbf{p}_{5,6,7,8}) = 0$.
9. Máx. en $\mathbf{p}_{1,2} = (\pm 1, 1)$, $f(\mathbf{p}_{1,2}) = 1$. Mín. en $\mathbf{p}_{3,4} = (\pm 1, -1)$, $f(\mathbf{p}_{3,4}) = -1$.
10. Máx. en $\mathbf{p}_{1,2} = (\pm 1, 1)$, $f(\mathbf{p}_{1,2}) = 9$. Mín. en $\mathbf{p}_3 = (0, -\sqrt{5}/2)$, $f(\mathbf{p}_3) = -4\sqrt{5}$.
11. Máx. en $\mathbf{p} = (3, 0)$, $f(\mathbf{p}) = 27$. Mín. en $(1, 2)$, $f(\mathbf{q}) = 15$.
12. Mín. en $\mathbf{p} = (7/12, -5/12)$, $f(\mathbf{p}) = 23/24$.
13. Máx. en $\mathbf{p} = (-4/3, 10/3)$, $f(\mathbf{p}) = 32/27$. Mín. en $\mathbf{q} = (0, 2)$, $f(\mathbf{q}) = 0$.
14. Máx. en $\mathbf{p} = (-15, 20)$, $f(\mathbf{p}) = 2375$. Mín. en $\mathbf{q} = (5/3, 10/3)$, $f(\mathbf{q}) = \dfrac{1625}{27}$.
15. Máx. en $\mathbf{p} = (-3, 9)$, $f(\mathbf{p}) = 540$. Mín. en $\mathbf{q} = (1, 5)$, $f(\mathbf{q}) = 156$
16. Máx. en $\mathbf{p} = (2, 4, 4)$, $f(\mathbf{p}) = 20$. Mín. en $\mathbf{q} = (-2, -4, -4)$ $f(\mathbf{q}) = -20$.
17. Máx. en $\mathbf{p} = (-1, -20, 25)$, $f(\mathbf{p}) = 206$. Mín. en $\mathbf{q} = (1, 20, -25)$ $f(\mathbf{q}) = -206$.
18. Máx. en $\mathbf{p} = (2\sqrt{5}, -\sqrt{5}, -2\sqrt{5})$, $f(\mathbf{p}) = 5\sqrt{5}$. Mín. en $\mathbf{q} = (-2\sqrt{5}, \sqrt{5}, 2\sqrt{5})$, $f(\mathbf{q}) = -5\sqrt{5}$.
19. Máx. en $\mathbf{p} = (4, -24, 8)$, $f(\mathbf{p}) = 88$. Mín. en $\mathbf{q} = (-4, 24, -8)$, $f(\mathbf{q}) = -88$.
20. Máx. en $\mathbf{p} = (5, 0, 2)$, $f(\mathbf{p}) = 7$. Mín. en $\mathbf{q} = (-5, 0, -2)$ $f(\mathbf{q}) = -7$.
25. Los semiejes son $a = 4$ (distancia máxima al origen en los puntos $\mathbf{p}_{1,2} = (\pm 2\sqrt{2}, \pm 2\sqrt{2})$, y $b = 2$ (distancia mínima al origen en los puntos $\mathbf{p}_{3,4} = (\pm\sqrt{2}, \mp\sqrt{2})$.
26. Área $= 12\pi$. (Los semiejes son 4 y 3).
27. Los puntos más cercanos son $\mathbf{p}_{1,2} = (\pm 2/\sqrt{3}, \pm 2/\sqrt{3})$, a una distancia $d = \sqrt{8/3}$. Los más alejados son $\mathbf{p}_{3,4} = (\pm 2, \mp 2)$, a una distancia $d = \sqrt{8}$.
28. El más alejado es el punto $\mathbf{p} = (1 + \sqrt{2}, 1 + \sqrt{2})$ a una distancia $d = 2 + \sqrt{2}$. El más cercano es el punto $q = (1 - \sqrt{2}, 1 - \sqrt{2})$ a una distancia $d = 2 - \sqrt{2}$.
29. Distancia mínima $= 1$ en el punto $(-2/3, -1/3, -2/3)$; distancia máxima $= 7$ en el punto $(14/3, 7/3, 14/3)$.
30. Distancia mínima $= 1.693$ en el punto $(9/\sqrt{13}, 4/\sqrt{13})$; distancia máxima $= 6.7921$ en el punto $(-9/\sqrt{13}, -4/\sqrt{13})$.
31. Los dos puntos $(\pm 3/\sqrt{8}, \pm 1/\sqrt{8})$.
32. El punto mas cercano, a una distancia de $11 - \sqrt{21}/3$, es $\mathbf{p} = (2 - 2/\sqrt{21}, 3 - 8/\sqrt{21}, 1 - 1/\sqrt{21})$ y el mas alejado, a una distancia de $11 + \sqrt{21}/3$, es $\mathbf{q} = (2 + 2/\sqrt{21}, 3 + 8/\sqrt{21}, 1 + 1/\sqrt{21})$.
33. El punto $\mathbf{p} = (3/2, 1, 25/4)$, que se encuentra a una distancia de $35\sqrt{3}/12$.
34. Los puntos $p = (\sqrt{3597}/3597, 2\sqrt{3597}/3597, -60\sqrt{3597}/1199)$, a una distancia de 3.98879, y $q = (-\sqrt{3597}/3597, -2\sqrt{3597}/3597, 60\sqrt{3597}/1199)$ a una distancia de 1.99376.
35. En el punto $\mathbf{p} = (1, 3, \sqrt{6} + 2)$ se tiene el mínimo que vale 0, y en el punto $q = (1, 1, 2 - \sqrt{15}/2)$ se tiene el máximo que vale $4\sqrt{15}/5$.
36. Los puntos de la elipse en donde se tiene el máximo son $(1/\sqrt{2}, 1/\sqrt{2}, -2/\sqrt{2})$ y $(-1/\sqrt{2}, -1/\sqrt{2}, 2/\sqrt{2})$, el cual vale 3; los puntos de la elipse en donde se tiene el mínimo

son $(1/\sqrt{2}, -1/\sqrt{2}, 0)$ y $(-1/\sqrt{2}, 1/\sqrt{2}, 0)$, el cual vale 1. Entonces el semieje mayor es $\sqrt{3}$ y el semieje menor es 1.

37. El valor máximo es $\sqrt{6}/18$ cuando dos de los números son $-\sqrt{6}/6$, y el otro es $\sqrt{6}/3$. El valor mínimo es $-\sqrt{6}/18$, cuando dos de los números son $\sqrt{6}/6$, y el otro es $-\sqrt{6}/3$.

38. Distancia $= \sqrt{966}/14$, en el punto $(6/7, 11/14, -15/14)$.

39. Distancia $= \sqrt{258}/3$, en el punto $(-1/3, 1/3, -7/3)$.

40. Distancia $= \sqrt{930}/18$, en el punto $(-43/54, -11/54, 40/27)$.

41. Distancia $= 1$, en el punto $(0, 0, 0)$.

42. Distancia $= 4\sqrt{6}/3$, en el punto $(1/3, 1/3, 1/3)$.

43. Distancia $= 3$, entre los puntos $(2, 1, -7)$ de L_1 y $(3, -1, -5)$ de L_2.

44. Distancia $= 6\sqrt{6}$, entre los puntos $(-15, -5, -20)$ de L_1 y $(-9, -17, -14)$ de L_2.

45. Distancia $= 0$ entre los puntos $(-7, -7, -7)$ de ambas rectas.

46. Distancia $= 32/31$ entre los puntos $(-1/31, -1/31, -1/31)$ de L_1 y $(-5/31, 23/31, -21/31)$ de L_2.

49. El cubo.

50. El cubo.

51. El cubo.

52. El cubo.

53. La recta tangente debe ser trazada en el punto $(a/\sqrt{2}, b/\sqrt{2})$. El área del triángulo así formado es $A_{\text{mín}} = ab$.

56. Los puntos más alejados al origen son $\mathbf{p}_{1,2} = (\mp 0.9883185681, \pm 2.3080244246)$ a una distancia $d_{\text{máx}} = 2.5107270533$, y los más cercanos del origen son $\mathbf{p}_{3,4} = (\pm 1.3787853648, \pm 0.2386561623)$ a una distancia $d_{\text{mín}} = 1.3999287609$.

57. Los puntos más cercanos al origen son $\mathbf{p}_{1,2} = \left(\sqrt{\sqrt{e-1}-1}, \sqrt{\sqrt{e-1}-1}\right)$ a una distancia $d_{\text{mín}} = 0.2751602091$. Los puntos más alejados del origen son $\mathbf{p}_3 = (-1.3271627913, -2.2064246412)$, $\mathbf{p}_4 = (-2.2064246412, -1.3271627913)$ a una distancia $d_{\text{máx}} = 2.5748147063$.

58. El punto más cercano al origen es $\mathbf{p}_1 = (-0.6614107536, 2.0558251672)$ a una distancia $d_{\text{mín}} = 2.1596021168$. El punto más alejado del origen es $\mathbf{p}_2 = (-1.3694675764, 3.9338314094)$ a una distancia $d_{\text{máx}} = 4.1653896577$.

59. El punto más cercano al origen es $\mathbf{p}_1 = (1.3156359355, 2.2152338968)$ a una distancia $d_{\text{mín}} = 2.57646252297$. El punto más alejado del origen es $\mathbf{p}_2 = (2.927575446, 3.6158007145)$ a una distancia $d_{\text{máx}} = 4.6523878599$.

60. El punto más cercano al origen es $\mathbf{p}_1 = (0.9552371645, 1.6742874885)$ a una distancia $d_{\text{mín}} = 1.9276194216$. El punto más alejado del origen es $\mathbf{p}_2 = (2.9606605249, 4.4412196064)$ a una distancia $d_{\text{máx}} = 5.33759705635$.

61. Los puntos más cercanos al origen son $\mathbf{p}_{1,2} = (\pm 0.1239828153, \pm 0.9870815682)$ a una distancia $d_{\text{mín}} = 0.994837554965$. Los puntos más alejados del origen son $\mathbf{p}_{3,4} = (\mp 1.0877380882, \pm 0.663635712)$ a una distancia $d_{\text{máx}} = 1.27420034012$.

62. Los puntos más cercanos al origen son $\mathbf{p}_{1,2} = (\mp 0.0916861692, \pm 0.6546442858)$ a una distancia $d_{\text{mín}} = 0.66103365614$. Los puntos más alejados del origen son $\mathbf{p}_{3,4} = (\pm 1.9005850981, 0.2661863407)$ a una distancia $d_{\text{máx}} = 1.91913493092$.

63. El punto más cercano al origen es $\mathbf{p}_1 = (0.1840669756, 0.5457821672, -0.3909518065)$ a una distancia $d_{\text{mín}} = 0.6961337089$. El punto más alejado del origen es

$\mathbf{p}_2 = (2.120577774389, 3.818818116028, -0.451654350256)$ a una distancia $d_{\text{máx}} = 4.391379458968$.

64. La distancia mínima es 3.409021745487, que se alcanza en el punto $\mathbf{p}_1 = (0.0442901647, 0.9990187092, -1.0433088739)$ de la intersección de las superficies, y el punto $q_1 = (0.6973602624, 4.3026397376, -0.5131986878)$ de la recta. La distancia máxima es 5.046582971062, que se alcanza en el punto $\mathbf{p}_2 = (-0.1472675754, -0.9890966895, 1.1363642649)$ de la intersección de las superficies, y el punto $q_2 = (1.1675426088, 3.8324573912, 1.8377130442)$ de la recta.

65. La distancia mínima es 1.14493921, alcanzada en el punto $\mathbf{p} = (0.6645447099, 0.1529546074, -0.5115901037)$. La distancia máxima es 3.775987249, alcanzada en el punto $q = (-1.6645447099, 0.3470453925, 2.011590103)$.

66. La distancia mínima es 5.69974, alcanzada en el punto (del elipsoide) $\mathbf{p} = (0.814598245, 0.13294487, 0.5487085034)$. La distancia máxima es 7.668338327, alcanzada en el punto (del elipsoide) $q = (-0.807033138, -0.120913845, -0.565205448)$.

67. Los puntos $\mathbf{p}_1 = (1.97972585, 0.1420260158)$ de la curva $x^2 + 4y^2 - 4 = 0$ y $q_1 = (2.80603385, 0.379144166)$ de la curva $x^2 + y^2 + 4x + 2y - 20 = 0$ son los que se encuentran más cercanos, a una distancia $d_{\text{mín}} = 0.859656866$. Los puntos $\mathbf{p}_2 = (1.97972585, 0.1420260158)$ de la curva $x^2 + 4y^2 - 4 = 0$ y $q_2 = (-6.8060338502, -2.3791441661)$ de la curva $x^2 + y^2 + 4x + 2y - 20 = 0$ son los que se encuentran más alejados, a una distancia $d_{\text{máx}} = 9.1403431333$.

68. Máximo absoluto en $(1, 0)$ en donde f vale $11/6$. Mínimo absoluto en (0,1) en donde f vale 0.

69. Máximo absoluto en $(1, 1)$ en donde f vale 2. Mínimo absoluto en $(-1, -1)$ en donde f vale -2.

70. Máximo absoluto en $(1, 1)$ en donde f vale 8. Mínimo absoluto en $(0, 0)$ en donde f vale 0.

71. Máximo absoluto en $(\sqrt{2}/2, \sqrt{2}/2)$ en donde f vale $1 + \sqrt{2}$. Mínimo absoluto en (-1/2,-1/2), en donde f vale $-1/2$.

72. Máximo absoluto en $\mathbf{p} = (0, 0)$ en donde f vale 2. Mínimos absolutos en $q_{1,2,3,4} = (\pm\pi, \pm\pi)$ en donde f vale -2.

73. Máximo absoluto en $(\pi/2, 0)$ en donde f vale 2. Mínimos absolutos en $(-\pi/2, \pm\pi)$ en donde f vale -2.

74. Máximo absoluto en $(-2, -2, -2)$ en donde f vale 27. Mínimo absoluto en $(1, 1, 1)$ (punto crítico) en donde f vale 0.

Capítulo 4, Sección 6 (página 423)

1. $H_F(\pm 2, \mp 3/2) = \pm 132$.

2. $H_F(\pm 1, \pm 6) = \pm 152$.

3. $H_F(\pm 3, \mp 28/3) = \pm 3568$.

4. $H_F(\pm 2, \pm 15) = \mp 948$.

5. $H_F(\pm 1, \pm 2) = \pm 20$.

6. $H_F(\pm 1/\sqrt{2}, \pm 1/\sqrt{2}) = H_F(\mp 1/\sqrt{2}, \mp 1/\sqrt{2}) = 8$. $H_F(\pm 1, 0) = H_F(0, \pm 1) = -8$.

7. $H_F(\pm 2, \pm 4, \pm 4) = -240$, $\Delta_3(\pm 2, \pm 4, \pm 4) = \pm 768$.

8. $H_F(\pm 1, \pm 20, \mp 25) = -824$, $\Delta_3(\pm 1, \pm 20, \mp 25) = \mp 1620$.

9. $H_F(\pm 2\sqrt{5}, \mp\sqrt{5}, \mp 2\sqrt{5}) = -20$, $\Delta_3(\pm 2\sqrt{5}, \mp\sqrt{5}, \mp 2\sqrt{5}) = \pm 24\sqrt{5}$.

10. $H_F(\pm 4, \mp 24, \pm 8) = -176$, $\Delta_3(\pm 4, \mp 24, \pm 8) = \pm 1280$.

11. $H_F(\pm 5, 0, \pm 2) = -112$, $\Delta_3(\pm 5, 0, \pm 2) = \pm 160$.

Capítulo 5, Sección 1 (página 430)

1. **a.** $(1, 2, -1)$, **b.** $(5, 3, 19)$, **c.** $(0.5, -2, -1.5)$, **d.** 0, **e.** $(\sqrt{77}, 2\sqrt{77}, -\sqrt{77})$, **f.** $(-10, 2, -6)$, **g.** 16
2. $I_1 \cap I_2 \cap \ldots \cap I_k$
3. $\mathbb{R}$
4. $\{t \in \mathbb{R} | t \geq 0\}$
5. $\{t \in \mathbb{R} | t \geq 1\}$
6. $\{t \in \mathbb{R} | t > 0\}$
7. $\mathbb{R}$
8. $\mathbb{R}$
9. $\{t \in \mathbb{R} | 0 < t \leq 1\}$
10. $\{t \in \mathbb{R} | -1/4 \leq t \leq 1/4\}$
11. $\delta \leq \epsilon/\sqrt{2}$
13. Cualquier $\delta > 0$
14. $(2, 1)$
15. $(2, 0)$
16. $(1, 0)$
17. $(1/2, 2)$
18. $(2, 3, 4)$
19. $(2x, 3x^2, 4x^3)$
20. $(2/3, 1, 0)$
21. Discontinua en $t = 1$
22. Discontinua en $t = 0$
23. Discontinua en $t = 1$ y $t = -1$
24. Continua en $\mathbb{R}$
25. Continua en $\mathbb{R}$

Capítulo 5, Sección 2 (página 441)

1. Traza de $\mathbf{f} = \{(a, b)\}$
4. No es simple pues $\mathbf{f}(-1) = \mathbf{f}(1) = (2, 0)$
6. No es simple pues $\mathbf{f}(-1) = \mathbf{f}(1) = (0, 0)$
7. Es un camino cerrado simple
8. La rama derecha de la hipérbola $x^2 - y^2 = 1$

Capítulo 5, Sección 3 (página 456)

2. $(0, \pi \cos \pi^2/4)$

3. $(3/2, 0)$

4. $(0, 0, 0)$

5. $(-e^{-1/2}, -1, 4/3)$

6. Es diferenciable y regular

7. Es diferenciable; no es regular, pues $\mathbf{f}'(0) = (0, 0, 0)$

8. Es diferenciable y regular

9. No es diferenciable

10. Es diferenciable; no es regular, pues $\mathbf{f}'(0) = (0, 0)$

11. Cierto

12. Falso, por ejemplo $\mathbf{f}(t) = (t, t)$

13. Falso, por ejemplo $\mathbf{f}(t) = (\cos t, \operatorname{sen} t)$

14. Cierto

16. $t = -3/2$

17. **a.** $t = 3\pi/2$, **b.** $t = \pi/2$, **c.** $t = 0$ y $t = 2\pi$,**d.** $t = \pi$, **e.** $t = \arctan(-b/a)$, **f.** $t = \arctan(-b/a) + \pi$

19. La recta es $\mathbf{f}(t)$, el plano es $x + 3y + 2z = 14t_0 - 12$

20. La recta es $x = t - 1$, $y = 2 + 2t$, $z = 3$, el plano es $x + 2y = 3$

21. La recta es $x = 1$, $y = 3t$, $z = 3t$, el plano es $y + z = 0$

22. La recta es $x = 1 - 2t$, $y = 1 - t$, $z = t$, el plano es $z - 2x - y = -3$

23. La recta es $x = t$, $y = 1 + t$, $z = 5t$, el plano es $x + y + 5z = 1$

24. La recta es $x = -\pi - t$, $y = 1$, $z = -2t$, el plano es $z = 0$

25. La recta es $x = -2$, $y = 5$, $z = 10 - 5t$, el plano es $z = 10$

26. En los puntos $f(2) = (-2, 12, 14)$ y $f(-1) = (-2, 3, -4)$

27. En el punto $f(1/2) = (3/2, -1/2, 0)$.

Capítulo 5, Sección 4 (página 467)

1. Sí

2. Sí

3. No

4. No

5. No

6. $(2, 15/2)$

7. $\varphi(s) = 1 - s^2/2$, $\bar{\mathbf{f}}$ no es una reparametrización de $\mathbf{f}$

8. $\varphi(s) = 2\pi(1 - s^2)$, $\mathbf{g}$ no es una reparametrización de $\mathbf{f}$

9. $\bar{\mathbf{f}}: [0, 1] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = (\text{sen}(5s), \cos(5s), 5s)$

10. $\bar{\mathbf{f}}: [0, 5] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = (1 - s/5, 1 - s/5, 1 - s/5)$

11. $\bar{\mathbf{f}}: [0, 1] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = (6s + 2, 8s^3 + 3)$

12. $\bar{\mathbf{f}}: [-2, 2] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = (3\cos(s/2), 3\,\text{sen}(s/2))$

13. $\bar{\mathbf{f}}: [0, 1/4] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = (16s^2 - 28s + 12, 11 - 16s)$

14. $\bar{\mathbf{f}}: [0, 3] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = ((3 - s)e^{3-s}, e^{s-3}, 3 - s)$

15. $\bar{\mathbf{f}}: [1, 4 + \pi] \to \mathbb{R}^3$, $\bar{\mathbf{f}}(s) = \left(\dfrac{3s + \pi}{3 + \pi}, \dfrac{2(3s + \pi)}{3 + \pi}, \dfrac{3(3s + \pi)}{3 + \pi}\right)$

16. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, 0)$

17. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (0, -t)$

18. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (t, 2t)$

19. $\mathbf{f}: [\pi/2, \pi] \to \mathbb{R}^2$, $\mathbf{f}(t) = (\sqrt{3}\cos t, \sqrt{3}\,\text{sen}\, t)$

20. $\mathbf{f}: [0, 4] \to \mathbb{R}^2$, $\mathbf{f}(t) = \begin{cases} (1 - t, t) & \text{si } t \in [0, 1) \\ (1 - t, 2 - t) & \text{si } t \in [1, 2) \\ (t - 3, 2 - t) & \text{si } t \in [2, 3) \\ (t - 3, t - 4) & \text{si } t \in [3, 4] \end{cases}$

21. $\mathbf{f}: [0, 5] \to \mathbb{R}^2$, $\mathbf{f}(t) = \begin{cases} (t, t/3) & \text{si } t \in [0, 3) \\ (6 - t, 4t - 11) & \text{si } t \in [3, 4) \\ (10 - 2t, 25 - 5t) & \text{si } t \in [4, 5] \end{cases}$

22. $\mathbf{f}: [-2, 2] \to \mathbb{R}^2$, $\mathbf{f}(t) = \begin{cases} (t, -1 - t) & \text{si } t \in [-2, -1) \\ (t, 1 + t) & \text{si } t \in [-1, 0) \\ (t, 1 - t) & \text{si } t \in [0, 1) \\ (t, t - 1) & \text{si } t \in [1, 2] \end{cases}$

23. $\mathbf{f}: [-2, 2] \to \mathbb{R}^2$, $\mathbf{f}(t) = \begin{cases} (t, t^2 - 1) & \text{si } t \in [-2, -1) \cup [1, 2] \\ (t, 1 - t^2) & \text{si } t \in [-1, 1) \end{cases}$

24. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (t, 0, 0)$

25. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (0, -t, 0)$

26. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (0, 0, -t)$

27. $\mathbf{f}: [0, \infty) \to \mathbb{R}^3$, $\mathbf{f}(t) = (t/2, t, t/3)$

28. $\mathbf{f}: [0, \infty) \to \mathbb{R}^3$, $\mathbf{f}(t) = (-t, 8t/9, 7t/9)$

29. $\mathbf{f}: [\pi/2, 5\pi/2] \to \mathbb{R}^3$, $\mathbf{f}(t) = (2\cos t, 2\,\text{sen}\, t, 4)$

Capítulo 5, Sección 5 (página 476)

1. $3\sqrt{26}$
2. $\sqrt{26}-\sqrt{2}-\ln(\sqrt{2}-\sqrt{26}+2\sqrt{13}-1)+\ln 5$
3. $(13\sqrt{13}-8)/27$
4. $3\,\text{senh}(2/3)$
5. $\ln(2+\sqrt{3})$
6. $3\sqrt{26}$
7. 24
8. $4+2^{3/2}\ln(1+\sqrt{2})$
9. $\sqrt{2}\,\text{senh}^2 1$
10. $\sqrt{2}\,\text{senh}\,1$
12. **c.** $\sqrt{2}e^{-t_0}$
14. 0
15. $(\beta-\alpha)\left(\sum_{i=1}^{n}a_i^2\right)^{1/2}$
17. Tomando como $t=0$ el momento del despegue, se tiene: ZONA 1: entrada a las 13.6749 horas, salida a las 17.8192 horas, distancia recorrida = 5466.72 km. ZONA 2: entrada a las 36.0051 horas, salida a las 40.4315 horas, distancia recorrida = 5838.83 km. ZONA 3: entrada a las 52.1323 horas, salida a las 60.5113 horas, distancia recorrida = 11052.6 km. Si $k\leq 70/9$, la nave del capitán Marcello no entraría a la zona de ondas expansivas de la explosión.

Capítulo 5, Sección 6 (página 483)

1. $\mathbf{f}''(s)=-r^{-1}(\cos(s/r),\text{sen}(s/r)),\ \|\mathbf{f}''(s)\|=r^{-1}$
2. $\mathbf{f}''(s)=(s^2+1)^{3/2}(-s,1),\ \|\mathbf{f}''(s)\|=1/(s^2+1)$
3. $\mathbf{f}''(s)=-\dfrac{\alpha}{\alpha^2+\beta^2}\left(\cos\dfrac{s}{\sqrt{\alpha^2+\beta^2}},\text{sen}\dfrac{s}{\sqrt{\alpha^2+\beta^2}},0\right),\ \|\mathbf{f}''(s)\|=|\alpha|/(\alpha^2+\beta^2)$

Capítulo 5, Sección 7 (página 502)

1. $1/5$
2. $\sqrt{2}/3$
3. $2\sqrt{2}/27$
4. $\sqrt{14}/12$
5. $\sqrt{38}/18$
6. $k(t)=-2\sqrt{2}e^{3t}(1+e^{4t})^{-3/2}$
7. $k(t)=\dfrac{6}{|t|(4+9t^2)^{3/2}}$

8. $k(t) = -\dfrac{6}{|t|(4+9t^2)^{3/2}}$

9. $k(t) = -\dfrac{2(3t^2-3t-1)}{(9t^4+12t^3+8t^2-4t+1)^{3/2}}$

10. $k(t) = 2e^{3t}(1+e^{4t})^{-3/2}$;

11. $k(t) = -\dfrac{\sqrt{2}}{4\sqrt{1-\cos t}}$

12. $k(t) = \dfrac{ab}{((b^2-a^2)\cos^2 t + a^2)^{3/2}}$, $k_{\text{máx}} = ab^{-2}$, para $t = 0, \pi, 2\pi$, $k_{\text{mín}} = ba^{-2}$, para $t = \pi/2, 3\pi/2$.

13. $k(x) = \dfrac{2}{(1+4x^2)^{3/2}}$, $k_{\text{máx}} = 2$ para $x = 0$. No hay mínimo.

14. $k(x) = \dfrac{6x}{(1+9x^2)^{3/2}}$. Para $x = -(1125)^{1/4}/15$ hay un mínimo local que vale $k_{\text{mín}} = -5(20)^{1/4}/6$, para $x = (1125)^{1/4}/15$ hay un máximo local que vale $k_{\text{máx}} = 5(20)^{1/4}/6$

15. $k(x) = \dfrac{n(n-1)x^n|x|}{(n^2x^{2n}+x^2)^{3/2}}$

16. $x^2 + (y-1/2)^2 = 1/4$

17. $x^2 + y^2 = 1$

18. En $(1,1)$: $(x+4)^2 + (y-8/3)^2 = 250/9$. En $(-1,-1)$: $(x-4)^2 + (y+8/3)^2 = 250/9$.

19. $x^2 + (y-1/2)^2 = 1/4$

20. $x^2 + (y-1/2)^2 = 1/4$

22. $\mathbf{f}: \mathbb{R} \to \mathbb{R}^2$, $\mathbf{f}(t) = (-4t^3, 3t^2 + 1/2)$

23. $\mathbf{f}: \mathbb{R} - \{0\} \to \mathbb{R}^2$, $\mathbf{f}(t) = \left((t-9t^5)/2, (15t^4+1)/6t\right)$

Capítulo 5, Sección 9 (página 525)

Ejercicio	1	2	3	4	5
recta tangente	$x = t$ $y = 0$ $z = 3$	$x = 1$ $y = t$ $z = t$	$x = t+1$ $y = 1-t$ $z = t$	$x = t+1$ $y = t$ $z = 2t+1$	$x = 3t+1$ $y = 2t+1$ $z = t+1$
recta normal	$x = 0$ $y = 2t$ $z = 2t+3$	$x = 1-2t$ $y = 0$ $z = 0$	$x = 3t+1$ $y = 3t+1$ $z = 0$	$x = 1-3t$ $y = -9t$ $z = 6t+1$	$x = 18t+1$ $y = 1-16t$ $z = 1-22t$
recta binormal	$x = 0$ $y = -2t$ $z = 2t+3$	$x = 1$ $y = -t$ $z = t$	$x = 1-t$ $y = 1+t$ $z = 2t$	$x = 4t+1$ $y = -2t$ $z = 1-t$	$x = 1-2t$ $y = 6t+1$ $z = 1-6t$
plano osculador	$z - y = 3$	$z - y = 0$	$-x + y + 2z = 0$	$4x - 2y - z = 3$	$x - 3y + 3z = 1$
plano normal	$x = 0$	$y + z = 0$	$x - y + z = 0$	$x + y + 2z = 3$	$3x + 2y + z = 6$
plano rectificante	$y + z = 3$	$x = 1$	$x + y = 2$	$x + 3y - 2z = -1$	$9x - 8y - 11z = -10$

6. $\det \begin{bmatrix} x-c_1 & y-c_2 & z-c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = 0$

8. Plano osculador: $128x + 81y - 5z = 750$, plano normal: $-3x + 4y - 12z = -12$, plano rectificante: $952x - 1551y - 755z = -1600$

9. Plano osculador: $z = 1$, plano normal: $y = x$, plano rectificante: $x + y = 2$

10. Recta tangente: $x = 1 + t$, $y = 1 + t$, $z = 1 + 2t$ recta normal: $x = 1 - t$, $y = 1 - t$, $z = 1 + t$ recta binormal: $x = 1 + t$, $y = 1 - t$, $z = 1$

11. $32/3$

Capítulo 5, Sección 10 (página 534)

1. $x - 4y + 2z = -1$

2. $f(t) = c_1 \operatorname{sen} at + c_2 \cos at + c_3$, donde c_1, c_2 y c_3 son constantes arbitrarias.

3. El valor común de la curvatura y de la torsión es $\dfrac{1}{3(t^2+1)^2}$

Capítulo 5, Sección 11 (página 550)

1. $\mathbf{r}''(t) = (0, 0) = (0)\,\mathbf{T}(t) + (0)\mathbf{N}(t)$

2. $\mathbf{r}''(1) = (0, 6) = \dfrac{9\sqrt{10}}{5}\mathbf{T}(1) + \dfrac{3\sqrt{10}}{5}\mathbf{N}(1)$

3. $\mathbf{r}''(-1) = (0, -6) = -\dfrac{9\sqrt{10}}{5}\mathbf{T}(-1) + \dfrac{3\sqrt{10}}{5}\mathbf{N}(-1)$

4. $\mathbf{r}''(0) = (-2, 0) = (0)\mathbf{T}(0) + (2)\mathbf{N}(0)$

5. $\mathbf{r}''(1) = (2, 6) = \dfrac{22\sqrt{13}}{13}\mathbf{T}(1) + \dfrac{6\sqrt{13}}{13}\mathbf{N}(1)$

6. $\mathbf{f}'(0) = (0, 2)$, $\mathbf{f}''(0) = (1, 0)$

9. $\sqrt{2}e^{\theta}/2$, la cual tiende a infinito cuando θ tiende a infinito.

10. $\dfrac{\theta^2+2}{(\theta^2+1)^{3/2}}$, la cual tiende a cero cuando θ tiende a infinito.

11. 4/3.

Capítulo 6, Sección 1 (página 560)

2. 20
3. 0
4. 16
5. 0
6. 21
7. 8
8. 20
9. 16
10. 29
11. −4
12. 35
13. 12
14. 128

Capítulo 6, Sección 2 (página 570)

1. 6
2. -4
3. -4
4. $-1/2$
5. 40
6. 48672
7. $(\operatorname{sen} 17 + \operatorname{sen} 26 - \operatorname{sen} 29 - \operatorname{sen} 14)/4$
8. $(2 \operatorname{sen} 1 - 2\cos 2 - 2\cos 1 - \operatorname{sen} 2 + 4)/4$
9. $e^5 - e^4 - e^2 + e$
10. $e(1 - \cos 1) + \cos 1 - 1$
11. $3 - e$
12. $1/80$

Capítulo 6, Sección 3 (página 587)

1. $\int_0^4 dx \int_0^{(4-y)/3} f(x, y)dy = \int_0^{4/3} dy \int_0^{4-3x} f(x, y)dx$
2. $\int_0^4 dx \int_0^{4x} f(x, y)dy = \int_0^{16} dy \int_{y/4}^4 f(x, y)dx$
3. $\int_1^5 dx \int_1^{(5x-1)/4} f(x, y)dy = \int_1^6 dy \int_{(4y+1)/5}^5 f(x, y)dx$
4. $\int_3^6 dy \int_{2y/3}^{8-(2y/3)} f(x, y)dx$
5. $\int_{-1}^1 dx \int_{-((1-x^2)/2)^{1/2}}^{((1-x^2)/2)^{1/2}} f(x, y)dy = \int_{-2^{-1/2}}^{2^{-1/2}} dy \int_{-(1-2y^2)^{1/2}}^{(1-2y^2)^{1/2}} f(x, y)\,dx$
6. $\int_0^1 dx \int_{x^2}^{\sqrt{x}} f(x, y)\,dy = \int_0^1 dy \int_{y^2}^{\sqrt{y}} f(x, y)\,dx$
7. $\int_0^2 dx \int_{2-(1-(x-1)^2)^{1/2}}^{2+(1-(x-1)^2)^{1/2}} f(x, y)\,dy = \int_1^3 dy \int_{1-(1-(y-2)^2)^{1/2}}^{1+(1-(y-2)^2)^{1/2}} f(x, y)\,dx$
8. $\int_{-1}^2 dx \int_{x^2}^{x+2} f(x, y)dy$
9. $\int_{-1}^1 dy \int_{y^2-1}^{1-y^2} f(x, y)dx$
10. $\int_{0.5(1-\sqrt{5})}^{0.5(1+\sqrt{5})} dy \int_{y^2+2}^{y+3} f(x, y)\,dx$

11. **a.** $\int_{-1}^{0} dx \int_{-x-1}^{x+1} f(x, y)dy + \int_{0}^{1} dx \int_{x-1}^{1-x} f(x, y)dy$

b. $\int_{-1}^{0} dy \int_{-y-1}^{y+1} f(x, y)dx + \int_{0}^{1} dy \int_{y-1}^{1-y} f(x, y)dx$

12. **a.** $\int_{-2}^{0} dx \int_{-3x/2}^{6+3x/2} f(x, y)dy + \int_{0}^{2} dx \int_{3x/2}^{6-3x/2} f(x, y)dy$

b. $\int_{0}^{3} dy \int_{-2y/3}^{2y/3} f(x, y)dx + \int_{3}^{6} dy \int_{2y/3-4}^{4-2y/3} f(x, y)dx$

13. **a.** $\int_{1}^{3} dx \int_{(x-1)/9}^{5(x-1)/2} f(x, y)dy + \int_{3}^{7} dx \int_{(x-1)/9}^{5} f(x, y)dy + \int_{7}^{10} dx \int_{(x-1)/9}^{(-4x+43)/3} f(x, y)dy$

b. $\int_{0}^{1} dy \int_{(2y+5)/5}^{9y+1} f(x, y)dx + \int_{1}^{5} dy \int_{(2y+5)/5}^{(43-3y)/4} f(x, y)dx$

14. **a.** $\int_{1}^{2} dx \int_{x/5}^{2x} f(x, y)dy + \int_{2}^{5} dx \int_{x/5}^{(x+18)/5} f(x, y)dy + \int_{5}^{7} dx \int_{2x-9}^{(x+18)/5} f(x, y)dy$

b. $\int_{0}^{1} dy \int_{y/2}^{5y} f(x, y)dx + \int_{1}^{4} dy \int_{y/2}^{(y+9)/2} f(x, y)dx + \int_{4}^{5} dy \int_{5y-18}^{(y+9)/2} f(x, y)dx$

15. **a.** $\int_{0}^{3} dx \int_{-x}^{4} f(x, y)dy + \int_{3}^{6} dx \int_{x-6}^{4} f(x, y)dy$

b. $\int_{-3}^{0} dy \int_{-y}^{y+6} f(x, y)dx + \int_{0}^{4} dy \int_{0}^{6} f(x, y)dx$

16. $\int_{0}^{1} dy \int_{y^2}^{\sqrt{y}} f(x, y)dx$

17. $\int_{0}^{1} dy \int_{2(1-y^2)^{1/2}}^{2(1+y^2)^{1/2}} f(x, y)\, dx$

18. $\int_{0}^{1} dy \int_{0}^{(1-y^2)^{1/2}} f(x, y)\, dx$

19. $\int_{0}^{1} dy \int_{y}^{10-9y} f(x, y)\, dx$

20. $\int_{0}^{4} dy \int_{\sqrt{y}-2}^{\sqrt{y}+2} f(x, y)\, dx$

21. $640/27$

22. 36

23. $\ln 2$

24. 0

25. $1/6$

26. 139/70

27. 15862/6

28. $(2/3)(e^2 - e - e^{-1} + e^{-2})$

29. $5e^3 - 15e + 6e^{-2} - 2$

30. $(9/4)\operatorname{sen} 2 - (3/8)\cos 2 - 2\operatorname{sen} 1 - 2\cos 1 + 5/8$

31. $(2/3)\pi ab$

32. 0

33. $-2/3$

34. 8/3

Capítulo 6, Sección 4 (página 606)

1. $\int_R\int f(x, y)\,dx\,dy = \int_\alpha^\beta du \int_\gamma^\delta f((bv - du)/(bc - da), (av - cu)/ad - bc))|1/(ad - bc)|\,dv$

2. $\int_R\int f(x, y)\,dx\,dy = \int_\alpha^\beta du \int_\gamma^\delta f((uv)^{1/2}, (u/v)^{1/2})|(1/2v)|\,dv$

3. $R' = \{(r, \theta)|0 \leq r \leq 2, 0 \leq \theta \leq \pi\}$

4. $R' = \{(r, \theta)|0 \leq r \leq 2, -\pi \leq \theta \leq 0\}$

5. $R' = \{(r, \theta)|1 \leq r \leq 3, -\pi/2 \leq \theta \leq \pi/2\}$

6. $R' = \{(r, \theta)|0 \leq r \leq 2\cos\theta, -\pi/2 \leq \theta \leq \pi/2\}$

7. $R' = \{(r, \theta)|0 \leq r \leq 4\operatorname{sen}\theta, 0 \leq \theta \leq \pi\}$

8. $R' = \{(r, \theta)|0 \leq r \leq -4\cos\theta, \pi/2 \leq \theta \leq 3\pi/2\}$

9. $R' = \{(r, \theta)|0 \leq r \leq -6\operatorname{sen}\theta, -\pi \leq \theta \leq 0\}$

10. $R' = \{(r, \theta)|2^{1/2} \leq r \leq 3^{1/2}, \pi/4 \leq \theta \leq \arctan 2\}$

11. $R' = \{(r, \theta)|1 \leq r \leq 3^{1/2}, 5\pi/4 \leq \theta \leq \pi + \arctan 3\}$

12. $R' = \{(r, \theta)|2^{1/2} \leq r \leq 2, 3\pi/4 \leq \theta \leq \pi\}$

13. 0

14. 50

15. 3/4

16. 8π

17. $\pi(e^9 - 1)$

18. $\ln(4/3)^\pi$

19. 24π

20. $7\pi^2/192$

21. $40960\pi/3$
22. $9/32$
23. 0
24. 0
25. $3^{-1/2}\ln(4/3)^{\pi}$

Capítulo 6, Sección 5 (página 621)

1. $\pi/2$
2. 15
3. $3\pi/2 - 2/3$
4. $9\ln^2 3 - 27\ln 3 + 18$
5. $2/3$
6. $\pi(1 - e^{-1})$
7. 12π
8. $8\pi/3$
9. $\pi(4\sqrt{6} - 22/3)$
10. 16π
11. $9\pi\sqrt{2}/4$
12. 2π
13. $\pi(4 - 2^{3/2})$
14. $7\pi/12$
15. $9/4$
16. $1/3$
17. $17^{3/2}/6$
18. $20\sqrt{5}/3$
19. 2
20. $3 - e$
21. $2^{3/2}$
22. $(3\pi - 2)/6$
23. $\pi/8$
24. $5\pi/32 + 1/12$
25. $1/60$
26. $\pi/4 + 1/2$

27. 13/216

28. $257\sqrt{3}/3456 + 1/48$

29. 5

30. ln 3

31. $a\sqrt{\alpha^2 - b^2} - ab \arctan \dfrac{\sqrt{\alpha^2 - b^2}}{b}$

32. Las dos áreas "laterales" son iguales a $2\pi/3 + 2^{-1/2} \ln(\sqrt{3} - \sqrt{2})$, El área "central" es igual a $8\pi/3 - \sqrt{2} \ln(\sqrt{3} - \sqrt{2})$

33. El área superior es $\dfrac{9}{200}(25\pi + 8) - \dfrac{9}{4} \arctan(3/4)$. El área inferior es $\dfrac{9}{4} \arctan(3/4) + \dfrac{9}{200}(25\pi - 8)$

34. $\dfrac{\pi a^2}{8} - \dfrac{ab}{4} + \dfrac{b^2 - a^2}{4} \arctan \dfrac{a}{b}$

36. El área "central" es igual a $2\pi ab - 4ab \arctan \dfrac{a}{b}$. Las dos áreas "laterales" son iguales a $2ab \arctan \dfrac{a}{b} - \dfrac{\pi ab}{2} = ab \arctan \dfrac{a^2 - b^2}{2ab}$.

37. $\dfrac{\pi b^2}{8} + \dfrac{a^2 - b^2}{4} \arctan \sqrt{-b/a} - \dfrac{\sqrt{-ab}(a + b)^2}{4(a - b)}$

38. $\dfrac{1}{6}a^2(2\pi + 3\sqrt{3})$

41. $\dfrac{3}{2}\pi a^2$

42. $\dfrac{9}{2}\pi$

43. $\dfrac{1}{8}\pi a^2$

44. $ka^2/3$

45. $\dfrac{\sqrt{3}a^2 k}{4}\left(\dfrac{\ln 3}{4} + \dfrac{1}{3}\right)$

46. Si los vértices del triángulo son $(0, 0)$, $(b, 0)$ y $(b/2, h)$, el centro de masa se encuentra en el punto $(b/2, h/3)$

47. El centro de masa se encuentra sobre la bisectriz del ángulo θ, a una distancia de $\dfrac{4}{3} R \dfrac{\operatorname{sen}(\theta/2)}{\theta}$ del centro

48. $(0, 3/5)$

49. $(5/2, 32/5)$

50. $(3\pi/16, 1/5)$

51. $ab^3/3$

52. $\pi R^4/8$

53. $4/9$

54. $5\pi R^4/4$

55. $bh^3/12$

56. $1/4$

57. $1/3$

58. c

59. $15/6$

60. $1 - e^{-1}$

61. $1/4$

62. $1/3$

Capítulo 6, Sección 6 (página 631)

1. a. $\int_{-1}^{1} dz \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} dy \int_{-\sqrt{1-y^2-z^2}}^{\sqrt{1-y^2-z^2}} f(x, y, z)dx$

b. $\int_{-1}^{1} dy \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dz \int_{-\sqrt{1-y^2-z^2}}^{\sqrt{1-y^2-z^2}} f(x, y, z)dx$

c. $\int_{-1}^{1} dz \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} dx \int_{-\sqrt{1-x^2-z^2}}^{\sqrt{1-x^2-z^2}} f(x, y, z)dy$

d. $\int_{-1}^{1} dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dz \int_{-\sqrt{1-x^2-z^2}}^{\sqrt{1-x^2-z^2}} f(x, y, z)dy$

e. $\int_{-1}^{1} dy \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} f(x, y, z)dz$

f. $\int_{-1}^{1} dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} f(x, y, z)dz$

2. a. $\int_{0}^{c} dz \int_{0}^{b(1-z/c)} dy \int_{0}^{a(1-y/b-z/c)} f(x, y, z)dx$

b. $\int_{0}^{b} dy \int_{0}^{c(1-y/b)} dz \int_{0}^{a(1-y/b-z/c)} f(x, y, z)dx$

c. $\int_{0}^{c} dz \int_{0}^{a(1-z/c)} dx \int_{0}^{b(1-x/a-z/c)} f(x, y, z)dy$

d. $\int_{0}^{a} dx \int_{0}^{c(1-x/a)} dz \int_{0}^{b(1-x/a-z/c)} f(x, y, z)dy$

e. $\int_{0}^{b} dy \int_{0}^{a(1-y/b)} dx \int_{0}^{c(1-x/a-y/b)} f(x, y, z)dz$

f. $\int_{0}^{a} dx \int_{0}^{b(1-x/a)} dy \int_{0}^{c(1-x/a-y/b)} f(x, y, z)dz$

3. $\int_{-1}^{1} dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy \int_{-1}^{1} f(x, y, z)\, dz,\quad \int_{-1}^{1} dy \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx \int_{-1}^{1} f(x, y, z)\, dz$

$\int_{-1}^{1} dx \int_{-1}^{1} dz \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} f(x, y, z)\, dy,\quad \int_{-1}^{1} dz \int_{-1}^{1} dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} f(x, y, z)\, dy$

$\int_{-1}^{1} dy \int_{-1}^{1} dz \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} f(x, y, z)\, dx,\quad \int_{-1}^{1} dz \int_{-1}^{1} dy \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} f(x, y, z)\, dx$

4. $\int_{-2}^{2} dy \int_{-2}^{2} dx \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} f(x, y, z)\, dz,\quad \int_{-2}^{2} dx \int_{-2}^{2} dy \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} f(x, y, z)\, dz$

$\int_{-2}^{2} dy \int_{-2}^{2} dz \int_{-\sqrt{4-z^2}}^{\sqrt{4-z^2}} f(x, y, z)\, dx,\quad \int_{-2}^{2} dz \int_{-2}^{2} dy \int_{-\sqrt{4-z^2}}^{\sqrt{4-z^2}} f(x, y, z)\, dx$

$\int_{-2}^{2} dx \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} dz \int_{-2}^{2} f(x, y, z)\, dy,\quad \int_{-2}^{2} dz \int_{-\sqrt{4-z^2}}^{\sqrt{4-z^2}} dx \int_{-2}^{2} f(x, y, z)\, dy$

5. (La respuesta no es única)

$$\int_{0}^{1} dx \int_{0}^{1-x} dy \int_{x+y-1}^{1-x-y} f(x, y, z) dz + \int_{0}^{1} dx \int_{x-1}^{0} dy \int_{x-y-1}^{1-x+y} f(x, y, z) dz$$
$$+ \int_{-1}^{0} dx \int_{0}^{1+x} dy \int_{-x+y-1}^{1+x-y} f(x, y, z) dz + \int_{-1}^{0} dx \int_{-x-1}^{0} dy \int_{-x-y-1}^{1+x+y} f(x, y, z) dz$$

6. $1/8$
7. $3/2$
8. $1/24$
9. $43/720$
10. $(\ln 2)/2 - 5/16$
11. $e^3 - 3e + 3e^{-1} - e^{-3}$
12. $7e/6 - 8/3$
13. $\pi^2 abc/4$

Capítulo 6, Sección 7 (página 644)

1. $\Omega = \{(r, \theta, z) | 0 \le r \le 1, 0 \le \theta \le \pi/2, -1 \le z \le 1\}$
2. $\Omega = \{(r, \theta, z) | 0 \le r \le 1, -\pi/2 \le \theta \le 0, -1 \le z \le 1\}$
3. $\Omega = \{(r, \theta, z) | 1 \le r \le 2, 0 \le \theta \le 2\pi, 0 \le z \le 1\}$
4. $\Omega = \{(r, \theta, z) | 1 \le r \le 2, \pi/4 \le \theta \le \pi/3, 0 \le z \le 2\}$
5. $\Omega = \{(r, \theta, z) | 0 \le r \le 2\cos\theta, -\pi/2 \le \theta \le \pi/2, 1 \le z \le 4\}$
6. $\Omega = \{(r, \theta, z) | 0 \le r \le 1, 0 \le \theta \le 2\pi, 0 \le z \le 2r^2\}$
7. $\Omega = \{(r, \theta, z) | 1 \le r \le 3, 0 \le \theta \le 2\pi, r^2 \le z \le 3r^2\}$
8. $\Omega = \{(r, \theta, \phi) | 0 \le r \le 1, 0 \le \theta \le 2\pi, 0 \le \phi \le \pi/2\}$

9. $\Omega = \{(r, \theta, \phi) | 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi, \pi/2 \leq \phi \leq \pi\}$

10. $\Omega = \{(r, \theta, \phi) | 0 \leq r \leq 2, \pi/2 \leq \theta \leq \pi, 0 \leq \phi \leq \pi/2\}$

11. $\Omega = \{(r, \theta, \phi) | 1 \leq r \leq 2, -\pi/2 \leq \theta \leq 0, \pi/2 \leq \phi \leq \pi\}$

12. $\Omega = \{(r, \theta, \phi) | 1 \leq r \leq 3, 0 \leq \theta \leq 2\pi, \pi/2 \leq \phi \leq \pi\}$

13. $\Omega = \{(r, \theta, \phi) | 0 \leq r \leq 2, 0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi/4\}$

14. $\Omega = \{(r, \theta, \phi) | 0 \leq r \leq 2, 0 \leq \theta \leq 2\pi, \pi/2 \leq \phi \leq 3\pi/4\}$

15. $\Omega = \{(r, \theta, \phi) | 1 \leq r \leq 3, \pi/4 \leq \theta \leq \pi/3, 0 \leq \phi \leq \pi\}$

16. $\pi/4$

17. 0

18. $(\pi/3)(\cos 1 - \cos 4)$

19. $32\pi/105$

20. $976\pi/3$

21. $4\pi/5$

22. $8\pi/15$

23. $2\pi - \pi^2/2$

24. -40π

25. $5\pi/24$

Capítulo 6, Sección 8 (página 654)

1. $16/3$

2. $7\pi/2$

3. π

4. 16π

5. $-4\pi/3$

6. $1/2$

7. $\pi(14/3 - 7\sqrt{2}/3)$

8. $abc\pi^2/4$

9. $abc\pi^2/(4\sqrt{2})$

10. $\pi a^2 bc/3$

11. El volumen es igual a $\frac{1}{3b}\pi a^2(b-a)^2$ y el área es igual a $\frac{\pi a}{b}(b-a)(b+3a)$.

12. 0

13. π

14. $(1/4, 1/4, 1/4)$

15. $(0, 0, 3/8)$

16. $(0, 3/8, 3/8)$

17. $(3/8, 3/8, 3/8)$

18. $(0, 0, 3/5)$

19. $(0, 0, 157/(20(9 \arctan(\sqrt{2}/2) + 3\sqrt{2} - 1)))$

20. $(0, 0, 1 + \sqrt{2})$

21. $(0, 0, 7/16)$

22. $I_x = \rho_0 abc(b^2 + c^2)/3$, $I_y = \rho_0 abc(a^2 + c^2)/3$, $I_z = \rho_0 abc(b^2 + a^2)/3$

23. $I_d = \rho_0 \pi^2 R^3/2$, $I_{\tan} = 128\rho_0 R^3/75$

24. $I_x = 8\rho_0 abc(b^2 + c^2)/3$, $I_y = 8\rho_0 abc(a^2 + c^2)/3$, $I_z = 8\rho_0 abc(b^2 + a^2)/3$

25. $1/2$

26. 0

27. $1/2$

28. $5/6$

29. $62/35$

30. 0

31. $3/4$

32. $1/2$

33. 0

Capítulo 6, Sección 9 (página 668)

2. $1/2$

3. $1/16$

4. 2

5. $1/36$

6. $1/24$

7. $1/384$

8. $1/48$

9. π^9

11. $1/6$

12. $5/21$

Respuestas a los ejercicios del capítulo 7,

Capítulo 7, Sección 2 (página 686)

10. grad $f(x, y) = (1, 1)$
11. grad $f(x, y) = (y, x)$
12. grad $f(x, y) = (2x, 2y)$
13. grad $f(x, y) = (4x, -2y)$
14. grad $f(x, y) = (2x, -4y)$
15. Otros campos que tienen la misma propiedad son, por ejemplo, $G_1(x, y, z) = (y, -x, 0)$, $G_2(x, y, z) = (-z, 0, x)$
16. $F_c(r, \theta, z) = r\mathbf{e}_r + z\mathbf{e}_z$, $F_e(r, \theta, \phi) = r\mathbf{e}_r$
17. $F_c(r, \theta, z) = r^2 \cos\theta \mathbf{e}_r + z\mathbf{e}_z$,
$F_e(r, \theta, \phi) = (r^2 \cos\theta \operatorname{sen}^3\phi + r\cos^2\phi)\mathbf{e}_r + (r^2\cos\theta\operatorname{sen}^2\phi\cos\phi - r\operatorname{sen}\phi\cos\phi)\mathbf{e}_\phi$
18. $F_c(r, \theta, z) = (r\operatorname{sen}\theta\cos\theta + r^2\operatorname{sen}^3\theta)\mathbf{e}_r + (r^2\operatorname{sen}^2\theta\cos\theta - r\operatorname{sen}^2\theta)\mathbf{e}_\theta + rz\operatorname{sen}\theta\cos\theta\mathbf{e}_z$
$F_e(r, \theta, \phi) =$
$(r\operatorname{sen}\theta\cos\theta\operatorname{sen}^2\phi + r^2\operatorname{sen}^3\theta\operatorname{sen}^3\phi + r^3\operatorname{sen}\theta\cos\theta\operatorname{sen}^2\phi\cos^2\phi)\mathbf{e}_r + (r^2\operatorname{sen}^2\theta\cos\theta\operatorname{sen}^2\phi - r\operatorname{sen}^2\theta\operatorname{sen}\phi)\mathbf{e}_\theta + (r\operatorname{sen}\theta\cos\theta\operatorname{sen}\phi\cos\phi + r^2\operatorname{sen}^3\theta\operatorname{sen}^2\phi\cos\phi - r^3\operatorname{sen}^2\theta\cos\theta\operatorname{sen}^2\phi\cos\phi)\mathbf{e}_\phi$
19. $F_c(r, \theta, z) = (r + r\operatorname{sen}\theta\cos\theta + z\operatorname{sen}\theta)\mathbf{e}_r + (-r\operatorname{sen}^2\theta + z\cos\theta)\mathbf{e}_\theta + z\mathbf{e}_z$
$F_e(r, \theta, \phi) = F_r(r, \theta, \phi)\mathbf{e}_r + F_\theta(r, \theta, \phi)\mathbf{e}_\theta + F_\phi(r, \theta, \phi)\mathbf{e}_\phi$,en donde
$F_r(r, \theta, \phi) = r\operatorname{sen}^2\phi(1 + \operatorname{sen}\theta\cos\theta) + r\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi + r\cos^2\phi$
$F_\theta(r, \theta, \phi) = -r\operatorname{sen}\theta\operatorname{sen}\phi(\cos\theta + \operatorname{sen}\theta) + r\cos\theta(\operatorname{sen}\theta\operatorname{sen}\phi + \cos\phi)$ $F_\phi(r, \theta, \phi) = r\cos\phi$
20. $F_c(r, \theta, z) = (r^2\operatorname{sen}\theta\cos^2\theta + rz\operatorname{sen}\theta\cos\theta)\mathbf{e}_r + (rz\cos^2\theta - r^2\operatorname{sen}^2\theta\cos\theta)\mathbf{e}_\theta + rz\operatorname{sen}\theta\mathbf{e}_z$
$F_e(r, \theta, \phi) = F_r(r, \theta, \phi)\mathbf{e}_r + F_\theta(r, \theta, \phi)\mathbf{e}_\theta + F_\phi(r, \theta, \phi)\mathbf{e}_\phi$, en donde
$F_r(r, \theta, \phi) = r^2\operatorname{sen}\theta\operatorname{sen}\phi(\cos^2\theta\operatorname{sen}^2\phi + \cos\theta\operatorname{sen}\phi\cos\phi + \cos^2\phi)$
$F_\theta(r, \theta, \phi) = r^2\cos\theta\operatorname{sen}\phi(-\operatorname{sen}^2\theta\operatorname{sen}\phi + \cos\theta\cos\phi)$
$F_\phi(r, \theta, \phi) = r^2\operatorname{sen}\theta\operatorname{sen}\phi\cos\phi(\cos^2\theta\operatorname{sen}\phi + \cos\theta\cos\phi - \operatorname{sen}\phi)$

Capítulo 7, Sección 3 (página 700)

1. 1
2. 0
3. 0
4. 0
5. 54/5
6. 2
7. $\dfrac{17}{24} - \dfrac{1}{4}\operatorname{sen} 2 + \dfrac{1}{32}\operatorname{sen} 4$
8. 0
9. 5/24
10. 52
11. −20
12. 5/3
13. **a.** 1/6, **b.** 0, **c.** 3/10, **d.** −1/2, **e.** 1/2
14. **a.** 4, **b.** 4, **c.** 4, **d.** 4, **e.** 4
15. **a.** 0, **b.** 0
16. **a.** 0, **b.** 0
17. **a.** 2, **b.** 2, **c.** 2
18. 0
19. **t1.** i, **t2.** $68 + 84i$, **t3.** 0, **t4.** 0, **t5.** $6 + 2i$, **u1.** $(1 + i)/3$,

u2. $(1+i)/3$, **u3.** $(1+i)/3$, **v11.** $-\pi i/2$, **v12.** $-\pi i/2$, **v13.** $-\pi i/2$, **v21.** $-i$, **v22.** $-1/6-4i/3$, **v23.** $-1-3i$, w1) $2\pi i$, **w2.** $2\pi i$, **w3.** 0, **x1.** $2\pi i$, **x2.** $4i$, **x3.** 5i

Capítulo 7, Sección 4 (página 722)

1. **b.** 4, **c.** $f(x, y) = x^2y + x + 2y^2$
2. **b.** 6, **c.** $f(x, y) = 3xy^3 + xy + 2x$
3. **b.** 9, **c.** $f(x, y) = x^2y^2 + xy + 5x + 2y$
4. **b.** 10, **c.** $f(x, y) = x^2y^3 + xy + x + 7y$
5. **b.** 4, **c.** $f(x, y) = x^3 + x^2y^2 + xy + y^3$
6. **b.** 9, **c.** $f(x, y) = 3x^5 + 2x^3y^3 + xy + 3y^5$
7. **b.** $e^2 + 2$, **c.** $f(x, y) = y^2e^{x+y} + x + y$
8. **b.** $3 + \operatorname{sen}^2 1$, **c.** $f(x, y) = 2x^2 + x\operatorname{sen}^2 y + y$
9. **b.** $2 \operatorname{sen} 1 + 2$, **c.** $f(x, y) = y \operatorname{sen} x + x \operatorname{sen} y + x + y$
10. **b.** $2 + e \operatorname{sen} 1$, **c.** $f(x, y) = e^x \operatorname{sen} y + x + y$
16. **b.** $\pi + e + e^{\pi}$, **c.** $f(x, y, z) = \pi x + ey + e^{\pi}z$
17. **b.** 1 , **c.** $f(x, y, z) = x^2y^3z$
18. **b.** e^3, **c.** $f(x, y, z) = xe^{y+2z}$
19. **b.** 5, **c.** $f(x, y, z) = 3xy + 2xz^2$
20. **b.** 4 , **c.** $f(x, y, z) = xz^2 + x + 2yz$
21. A lo largo de λ es igual a e; a lo largo de μ es igual a 0
22. La integral es igual a 0; la función potencial es $f(x, y, z, u) = 2x^2yu + 3xyz + z^3u^2$

Capítulo 7, Sección 5 (página 739)

Ejercicio	¿convexo?	¿conexo?	¿conexo por caminos?	¿simplemente conexo?
1	sí	sí	sí	sí
2	no	no	no	no
3	no	sí	sí	sí
4	no	no	no	no
5	no	sí	sí	sí
6	sí	sí	sí	sí
7	no	sí	sí	no
8	no	sí	sí	no
9	sí	sí	sí	sí
10	no	no	no	no

Capítulo 7, Sección 6 (página 751)

1. $x^2y + x + 2y^2 = c$
2. $xy + x + y = c$
3. $x^6y + 7x + 9y = c$
4. $x^2y^2 + xy + 5x + 2y = c$
5. $2x^2y^3 + 5xy^4 + 4x + 8y = c$
6. $x^2y^3 + xy + x + 7y = c$
7. $3xy^3 + xy + 2x = c$
8. $12x^2y^3 + x^2 + y^2 = c$
9. $x^2y^3 + x^2 + y^3 = c$
10. $2x^2 + x\operatorname{sen}^2 y + y = c$

11. $6xy^7 + 5xy + y = c$
12. $2xy^2 + 10xy + 2x + 3y = c$
13. $2x^3 + 10x^2y + 20xy + y^2 = c$
14. $y^2e^{x+y} + x + y = c$
15. $x\,\text{sen}(x + y) + y + x = c$
16. $y\,\text{sen}(x + y) + x\cos(x - y) = c$
17. $\mu(x) = e^x, (x - 2y)ye^x = c$
18. $\mu(y) = y, xy^2(x^2 + y) = c$
19. $\mu(x) = x^3, 3x^5y + x^4y^3 + 12x^4y^2 = c$
20. $\mu(xy) = xy, (3x + 7y + 6)x^2y^2 = c$
21. $\mu(x^2 + y^2) = e^{x^2+y^2}, (6x^2 + y)e^{x^2+y^2} = c$
22. $\mu(x^5 + y^5) = (x^5 + y^5)^2, (3y + 1)(x^5 + y^5)^3 = c$
23. $\mu(x + y^2) = (x + y^2)^{-1}, (3x + 4y)\ln(x + y^2) - 3y^2 = c$
24. $y = (\text{sen}\,x + \cos x)/2 + ce^{-x}$
25. $y = e^{3x}/5 + ce^{-2x}$
26. $y = x^2/3 + c/x$
27. $y = 0.5x + 0.5(1 + x^2)^{-1/2}(c + \ln(x + (1 + x^2)^{1/2})$
28. $y = 1/6 - (3/26)\cos 2x - (1/13)\,\text{sen}\,2x + ce^{-3x}$
30. **a.** $y^{-1} = 1 + ce^x$, **b.** $y^{-1} = e^x/2 + ce^{3x}$, **c.** $y^{-1} = \left(1 + ce^{x^2}\right)^{1/2}$
31. **d1** $x = c(u - 1)^3 - (3u - 2)^{5/3-3/4}$, en donde $u = y/x$,
d2 $6x^2 + 2xy + 7y^2 = c$, **d3** $x + (x^2 + y^2)^{1/2} = c$,
d4 $(y - x)^2(2y + x) = c$

Capítulo 7, Sección 7 (página 768)

1. $2\sqrt{2}$
2. $10\sqrt{10}$
3. π
4. 0
5. 2π
6. $-7\sqrt{5}/16 - 17\ln(2 + \sqrt{5})/32$
7. $\text{senh}\,2 + (1/3)\,\text{senh}^3\,2$
8. $78\sqrt{2}$
9. $2\pi^2\sqrt{2}$
10. $\dfrac{\frac{\pi}{4}(1 + \pi^2)(2 + \pi^2)^{1/2} + \frac{1}{2}\ln\sqrt{2}}{\pi + (2 + \pi^2)^{1/2} + (2 + \pi^2)^{3/2} - 2^{3/2}}$
11. **a.** k, **b.** k, **c.** k, **d.** k, **e.** $2k$, **f.** $4k$, **g.** $2k$
12. **a.** $k_1 + k_2$, **b.** $12k_1 - 4k_2$, **c.** $8k_1 - 24k_2$
13. $(1/3)(1 + e^2)^{3/2} - 2^{3/2}$
14. $2\sqrt{2}/3$
15. $(0, 0)$
16. $(0, 1/2)$
17. $(0, (4\,\text{senh}\,1)^{-1}(2 + \text{senh}\,2))$
18. $\left(1, \dfrac{h(1 + h^2)^{1/2}}{2(1 + (1 + h^2)^{1/2})}\right)$
19. $\pi R/4$

21. $4\pi r\sqrt{R^2 - r^2}$
22. $4a^2$
23. **a.** $2a^2$, **b.** $0.5\pi a^3$
24. aL

Capítulo 7, Sección 8 (página 778)

1. 0
2. 0
3. **a.** $4\sqrt{2}$, **b.** $4\sqrt{2}$, **c.** $4\sqrt{2}$
4. **a.** 2, **b.** 2, **c.** 2;
5. **a.** 3/2, **b.** 3/2;
6. 4

Capítulo 7, Sección 9 (página 796)

6. -8π
7. $-75\pi/2$
8. $2a^2$
12. $-3/5$
13. πab
14. $2a^2$
15. $(3/8)\pi a^2$
19. Sí, por ejemplo, un círculo de radio 2.

Capítulo 7, Sección 10 (página 806)

1. 2
2. 2/3
3. 0
4. $y^2 - x^2$
6. Sí es conservativo

Capítulo 7, Sección 11 (página 818)

5. 0
6. 16/3
7. 0
8. 2
9. $20\pi/\sqrt{3}$
10. $45\pi/2$
11. 7/2
12. 5
13. $3\pi/2 - 1/3$
14. $2z + 1/r$
15. $4r^2\theta + 2rz$
16. 3/2
17. $4 + \sqrt{2} + \pi/4$
18. $3(1 + \sqrt{2})/2$
19. $3\cos\theta\cos\phi + \cot\phi - \operatorname{sen}\phi$
20. $4r + 2r^{-1}\cos\phi - r^{-1}\cos\theta\csc\phi + 2r^{-1}\cos\phi$

Respuestas a los ejercicios del capítulo 8,

Capítulo 8, Sección 1 (página 832)

1. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 3/4\}$,
$f(u, v) = (u, v, (1 - u^2 - v^2)^{1/2})$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (\frac{\sqrt{3}}{2}\cos t, \frac{\sqrt{3}}{2}\operatorname{sen} t, \frac{1}{2})$
c. $\operatorname{Int}(K) = \{(x, y, z)|x^2 + y^2 + z^2 = 1, \frac{1}{2} < z \leq 1\}$
2. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 3/2\}$,
$f(u, v) = (\frac{1}{\sqrt{2}}(2 - u^2 - v^2)^{1/2}, u, v)$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (\frac{1}{2}, (3/2)^{1/2}\cos t, (3/2)^{1/2}\operatorname{sen} t)$
c. $\operatorname{Int}(K) = \{(x, y, z)|2x^2 + y^2 + z^2 = 2, \frac{1}{2} < x \leq 1\}$
3. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 8/3\}$, $f(u, v) = (u, -\frac{1}{\sqrt{3}}(3 - u^2 - 3v^2)^{1/2}, v)$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = ((8/3)^{1/2}\cos t, -1/3, (8/3)^{1/2}\operatorname{sen} t)$
c. $\operatorname{int}(K) = \{(x, y, z)|x^2 + 3y^2 + 3z^2 = 3, -1 \leq y < -1/3\}$
4. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|3u^2 + v^2 \leq 4\}$, $f(u, v) = \{u, v, 3u^2 + v^2\}$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = ((2/\sqrt{3})\cos t, 2\operatorname{sen} t, 4)$
c. $\operatorname{int}(K) = \{(x, y, z)|z = 3x^2 + y^2, 3x^2 + y^2 < 4\}$
5. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 4u\}$,
$f(u, v) = (u, v, u^2 + v^2)$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (2\cos t + 2, 2\operatorname{sen} t, 8(\cos t + 1))$
c. $\operatorname{int}(K) = \{(x, y, z)|z = x^2 + y^2, x^2 + y^2 < 4x\}$
6. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 1\}$,
$f(u, v) = (u, v, e^{-(u^2+v^2)})$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (\cos t, \operatorname{sen} t, e^{-1})$
c. $\operatorname{int}(K) = \{(x, y, z)|z = e^{-(x^2+y^2)}, x^2 + y^2 < 1\}$
7. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 3v\}$,
$f(u, v) = (u, v, u^2 + v^2 + 3u - 8v + 1)$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (\frac{3}{2}\cos t, \frac{3}{2}\operatorname{sen} t + \frac{3}{2}, \frac{9}{2}\cos t - \frac{15}{2}\operatorname{sen} t - \frac{13}{2})$
c. $\operatorname{int}(K) = \{(x, y, z)|z = x^2 + y^2 + 3x - 8y + 1, x^2 + y^2 < 3y\}$
8. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|0 \leq u \leq 1, 0 \leq v \leq u\}$, $f(u, v) = (u, v, -u^2 - v^2)$
b. $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, $\lambda_1, \lambda_2, \lambda_3\colon [0, 1] \to \mathbb{R}^3$, $\lambda_1(t) = (t, 0, -t^2)$, $\lambda_2(t) = (1, t, -1 - t^2)$,
$\lambda_3(t) = (1 - t, 1 - t, -2(1 - t)^2)$
c. $\operatorname{int}(K) = \{(x, y, z)|z = -x^2 - y^2, 0 < x < 1, 0 < y < x\}$
9. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|u^2 + v^2 \leq 4\}$, $f(u, v) = (1 + u^2 + v^2, u, v)$
b. $\lambda\colon [0, 2\pi] \to \mathbb{R}^3$, $\lambda(t) = (5, 2\cos t, 2\operatorname{sen} t)$
c. $\operatorname{int}(K) = \{(x, y, z)|x - y^2 - z^2 = 1, y^2 + z^2 < 4\}$
10. **a.** $f\colon S \to \mathbb{R}^3$, $S = \{(u, v)|0 \leq u \leq 1, 0 \leq v \leq 1 - u\}$ $f(u, v) = (u, u^2 + v^2 + 2u + 2v + 8, v)$
b. $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, $\lambda_1, \lambda_2, \lambda_3\colon [0, 1] \to \mathbb{R}^3$, $\lambda_1(t) = (t, t^2 + 2t + 8, 0)$, $\lambda_2(t) = (1 - t, 2t^2 + 11, t)$,
$\lambda_3(t) = (0, t^2 - 6t + 13, 1 - t)$
c. $\operatorname{int}(K) = \{(x, y, z)|y = x^2 + z^2 + 2x + 4z + 8, 0 < x < 1, 0 < z < 1 - x\}$

Capítulo 8, Sección 2 (página 838)

1. Sí se puede obtener una reparametrización,
2. Sí se puede obtener una reparametrización,
3. Sí se puede obtener una reparametrización
4. No se puede obtener una reparametrización (la función φ no es sobreyectiva).
5. Sí se puede obtener una reparametrización,
6. Sí se puede obtener una reparametrización,
7. Sí se puede obtener una reparametrización
8. No se puede obtener, en general, una reparametrización; a menos que $a = -\gamma$, $b = \gamma$, $c = -\gamma$, $d = \gamma$, para algún $\gamma > 0$ (es decir, que S sea un cuadrado con centro en el origen)
9. No se puede obtener, en general, una reparametrización; a menos que $a = -\gamma$, $b = \gamma$, $c = -\gamma$, $d = \gamma$, para algún $\gamma > 0$ (es decir, que S sea un cuadrado con centro en el origen)
10. Sí se puede obtener una reparametrización
11. No se puede obtener una reparametrización (la función φ no es biyectiva).
12. Sí se puede obtener una reparametrización.

Capítulo 8, Sección 3 (página 846)

Ejercicio	Espacio tangente	Plano tangente
3	$z = 2x + 8y$	$z = 2x + 8y - 5$
4	$y = 6(x + z)$	$6x - y + 6z = 19$
5	$4x + 20y + 3z = 0$	$4x + 20y + 3z = -16$
6	$x + 2y + 3z = 0$	$x + 2y + 3z = 14$
7	$x = 0$	$x = 1$
8	$z = 6x$	$z = 6x - 3$
9	$y = 6z$	$y = 6z - 8$
10	$z = -4e^{-2}x$	$z = e^{-2}(5 - 4x)$

11. **a.** Espacio tangente: $z = 0$; Plano tangente: $z = \varphi(x_0, y_0)$
b. Espacio tangente: $y = 0$; Plano tangente: $y = y_0$
c. Espacio tangente: $x = 0$; Plano tangente: $x = x_0$

Capítulo 8, Sección 4 (página 856)

1. Sean $S_1 = \{(u, v)|(u - u_0)^2 + (v - v_0)^2 \leq 1\}$, $S_2 = \{(u, v)|u^2 + v^2 \leq 1\}$, tomando u_0 y v_0 tales que $S_1 \cap S_2 = \varnothing$. Una parametrización de K está dada por $f\colon S_1 \cup S_2 \to \mathbb{R}^3$,

$$f(u, v) = \begin{cases} (u - u_0, v - v_0, (u - u_0)^2 + (v - v_0) - 1) & \text{si } (u, v) \in S_1 \\ (u, v, (1 - u^2 - v^2)^{1/2}) & \text{si } (u, v) \in S_2 \end{cases}$$

Plano tangente a K en $q_1\colon z = -1$; Plano tangente a K en $q_2\colon z = 1$. No es posible trazar un plano tangente a K en $(1, 0, 0)$.

2. Sean $S_1 = \{(u, v)|(u - u_0)^2 + (v - v_0)^2 \leq 1\}$, $S_2 = [0, 2\pi] \times [-1, e^{-1}]$, tomando u_0 y v_0 tales que $S_1 \cap S_2 = \varnothing$. Una parametrización de K está dada por $f\colon S_1 \cup S_2 \to \mathbb{R}^3$

$$f(u, v) = \begin{cases} (u - u_0, v - v_0, \exp[-(u - u_0)^2 - (v - v_0)^2]) & \text{si } (u, v) \in S_1 \\ (\cos u, \operatorname{sen} u, v) & \text{si } (u, v) \in S_2 \end{cases}$$

Plano tangente a K en $q_1\colon z = 1$; Plano tangente a K en $q_2\colon y = -x + \sqrt{2}$. No es posible trazar un plano tangente a K en $(1, 0, e^{-1})$.

3. Sean $S_1 = [0, 2\pi] \times [-2, 0]$, $S_2 = \{(u, v) | 1 \leq (u - u_0)^2 + (v - v_0)^2 \leq 4\}$, $S_3 = \{(u, v) | (u - u_1)^2 + (v - v_1)^2 \leq 1\}$ tomando u_0, v_0, u_1, v_1 tales que $S_1 \cap S_2 \cap S_3 = \varnothing$. Una parametrización de K está dada por $f: S_1 \cup S_2 \cup S_3 \to \mathbb{R}^3$,

$$f(u, v) = \begin{cases} (\cos u, \operatorname{sen} u, v) & \text{si } (u, v) \in S_1 \\ (u - u_0, v - v_0, 4 - (u - u_0)^2 - (v - v_0)^2) & \text{si } (u, v) \in S_2 \\ (u - u_1, v - v_1, 3 + (1 - (u - u_1)^2 - (v - v_1)^2)^{1/2}) & \text{si } (u, v) \in S_3 \end{cases}$$

Capítulo 8, Sección 6 (página 871)

1. $\frac{1}{2}\sqrt{a^2b^2 + a^2c^2 + b^2c^2}$
2. $4a^2 \operatorname{arcsen} \frac{b}{a}$
3. $2a^2(\pi - 2)$
4. $\sqrt{2}\pi$
5. $16a^2$
6. $4\pi(3 + 2\sqrt{3})a^2$
8. $\frac{2\pi}{3}((1 + a^2)^{3/2} - 1)$

Respuestas a los ejercicios del capítulo 9,

Capítulo 9, Sección 1 (página 891)

2. $\pi/2$
3. 0
4. $\pi/2$
5. $-3\pi/2$
6. 0
7. $-\pi$
8. $\sqrt{3}/12$
9. $\sqrt{38}/21600$
10. $8[(b - a)(d - c) + (b - a)(f - e) + (d - c)(f - e)]$
11. $5\sqrt{3}$
12. 0
13. La integral vale $2\pi(b - a)$. El valor medio es 1
14. La integral vale $4\pi c^4$. El valor medio es c^2
16. 0
17. 0
19. $1/60$
20. $3/2$
21. $1/2$
22. πRH^2
23. $(0, 0, 1/2)$

Capítulo 9, Sección 2 (página 904)

2. 2
3. 0
4. $1/2$

5. 4
6. 0
7. 0
8. 0
9. 0
10. 1/8
11. 0
12. −5/432

Capítulo 9, Sección 3 (página 914)

6. $12\pi/5$
10. **a.** $3\pi^2/4$, **b.** 4π
11. **a.** $4\pi/3$, **b.** 2π
12. **a.** $\pi/4$, **b.** $2\pi/3$
13. **a.** 4π, **b.** π

Capítulo 9, Sección 4 (página 925)

1. $(0, 0, 0)$
2. $(0, 0, 0)$
3. $(xz - 3, -yz, 0)$
4. $(2z, -3z, 1 - x^2)$
5. $(0, 0, -xe^x)$
6. $\frac{1}{r}\mathbf{e}_z$
7. $\frac{\theta}{r}\mathbf{e}_z$
8. $(1 + r^{-1}) \operatorname{sen} \theta \mathbf{e}_z$
9. $r^{-1} \cos\theta(\mathbf{e}_r + \mathbf{e}_z)$
10. $r^{-1} \operatorname{sen}(2t)\mathbf{e}_r + (r^{-1} \operatorname{sen}(2t) + 4r^2)\mathbf{e}_z$
11. $r^{-1} \cot\phi \mathbf{e}_r - r^{-1}\mathbf{e}_\phi + r^{-1}\mathbf{e}_\theta$
12. $r^{-1}(\phi \cot\phi - \csc\phi + 1)\mathbf{e}_r - r^{-1}\phi\mathbf{e}_\phi + r^{-1}\theta\mathbf{e}_\theta$
13. $\csc\phi(2\cos^2\phi - 1)\mathbf{e}_r + (\csc\phi\cos\theta - 2\cos\phi)\mathbf{e}_\phi$
14. $(\cos\theta\cos^2\phi\csc\phi - r^{-1}\cos\theta\cot\phi - \cos\theta \operatorname{sen}\phi)\mathbf{e}_r - 2\cos\theta\cos\phi\mathbf{e}_\phi + r^{-1} \operatorname{sen}\theta\cos\phi\mathbf{e}_\theta$
15. $-r^{-1} \operatorname{sen}\theta\mathbf{e}_\phi + (4 \operatorname{sen}\phi - r^{-1}\cos\theta\cos\phi)\mathbf{e}_\theta$

Respuestas a los ejercicios del capítulo 10,

Capítulo 10, Sección 1 (página 956)

1. Es una 0−forma en $\mathbb{R}$
2. Es una 0−forma en $\mathbb{R}^3$
3. Es una 1−forma en $\mathbb{R}^2$
4. Es una 2−forma en $\mathbb{R}^2$
5. Es una 3-forma en $\mathbb{R}^4$
6. Es una 5−forma en $\mathbb{R}^5$
7. Es una 3-forma definida en el conjunto abierto U de $\mathbb{R}^6$ dado por

$$U = \{(x_1, x_2, \ldots, x_6) \in \mathbb{R}^6 | x_1 > 0, x_2 > 0\}$$

8. Es una 3−forma en $\mathbb{R}^4$
9. Es una 4−forma en $\mathbb{R}^7$
10. Es una 4−forma en $\mathbb{R}^{12}$
11. $\omega\eta = x_1 \operatorname{sen} x_1$
12. $\omega\eta = -(x_1^2 + x_2^2)dx_1dx_2$
13. $\omega\eta = 0$
14. $\omega\eta = x_1x_2x_3x_4x_7dx_1dx_2dx_3dx_4dx_5dx_6 - x_2x_4x_7(3x_1 - x_2)dx_1dx_2dx_3dx_4dx_6dx_7$
15. $\omega\eta = 0$
16. $\omega\eta = \eta\omega = x_1(x_1 + x_2)dx_1dx_3 - x_3(x_1 + x_2)dx_2dx_3$
17. $\omega\eta = \eta\omega = 18x_1x_2^2dx_1dx_2dx_3 + 3x_1x_2dx_1dx_3dx_4 - 2x_1x_4dx_1dx_2dx_4$
18. $\omega\eta = \eta\omega = 2x_3dx_1dx_2dx_3dx_4dx_5 + x_1x_2x_3dx_1dx_2dx_3dx_4dx_5$
19. $\omega\eta = \eta\omega = -2dx_1dx_2dx_3dx_5dx_7 - dx_1dx_2dx_3dx_4dx_8 + dx_2dx_3dx_5dx_7dx_8$
20. $\omega\eta = \eta\omega =$
$4x_1x_5dx_1dx_2dx_3dx_5 - 2x_4dx_1dx_4dx_5dx_7 + 2x_7x_8dx_1dx_5dx_7dx_8 + 16dx_1dx_2dx_4dx_5 +$
$4x_7x_8dx_2dx_4dx_7dx_8 - 4x_2x_4x_8dx_1dx_2dx_4dx_8 + 4x_1dx_1dx_3dx_5dx_8 + x_1x_4dx_3dx_4dx_7dx_8 -$
$2x_1x_5x_7x_8dx_2dx_3dx_7dx_8 + 4x_1x_3x_7dx_1dx_2dx_5dx_7 + x_1x_2x_3x_4x_7x_8dx_1dx_2dx_7dx_8$
21. **a.** $x_1x_2(x_2x_3 + x_1x_3 - 1)dx_1 - (x_2^2x_3^2 - x_1)dx_2 + x_2x_3(x_3 - x_2)dx_3$
b. $x_1x_3(x_2^2 - (1 + 3x_2x_3))dx_1dx_3 - x_1x_2(3 + x_3(1 + 3x_2x_3))dx_1dx_2 + +x_3(x_2^2x_3 + 3)dx_2dx_3$
c. el negativo del resultado del inciso anterior

Capítulo 10, Sección 2 (página 969)

1. $d\omega = 0$
2. $d\omega = -x_1x_3dx_1dx_2 - x_1x_2dx_1dx_3$
3. $d\omega = 0$
4. $d\omega = -x_1^2dx_1dx_2dx_4$
5. $d\omega = \operatorname{sen} x_2dx_1dx_2dx_4 + x_5dx_2dx_3dx_5 + x_3dx_1dx_4dx_5 + x_4x_5dx_2dx_4dx_5 + +x_1dx_3dx_4dx_5$
6. $d\omega = x_4^3dx_1dx_2dx_3dx_4$
7. $d\omega = 0$
8. $d\omega = -x_2 \operatorname{sen} x_3 \operatorname{sen} x_4dx_2dx_3dx_4dx_5$
9. $d\omega = x_3x_6^3dx_2dx_3dx_5dx_6$
10. $d\omega = [x_1x_2x_3x_4 \operatorname{sen}(2x_1x_2) - x_3x_4\cos^2(x_1x_2)]dx_1dx_2dx_5dx_7 - x_2x_4\cos^2(x_1x_2)dx_1dx_3dx_5dx_7 - x_2x_3\cos^2(x_1x_2)dx_1dx_4dx_5dx_7 + [x_3x_6 \operatorname{sen}^2(x_1 + x_3x_6) + x_1x_3x_6 \operatorname{sen}(2x_1 + 2x_3x_6)]dx_1dx_3dx_6dx_7$

Capítulo 10, Sección 3 (página 978)

1. $\varphi^*\omega = y_1dy_1 + y_2dy_2$
2. $\varphi^*\omega = y_1y_2^2y_3^2dy_1 + (y_1^2y_2y_3^2 + y_2y_3^2)dy_2 + (y_1^2y_2^2y_3 + y_2^2y_3)dy_3$
3. $\varphi^*\omega = (y_1^2y_2 - y_1y_2)dy_1dy_2$
4. $\varphi^*\omega = 2y_1^3y_2^3y_3^3y_4dy_1dy_3dy_4 + 2y_1^3y_2y_3^4y_4dy_1dy_2dy_4 + 2y_1^3y_2y_3^3y_4^2dy_1dy_2dy_3$
5. $\varphi^*\omega = [y_1y_2y_3 \operatorname{sen}(y_1 + y_2 + y_3) + y_3 + y_2y_3\cos(y_1 - y_2 - y_3)]dy_1dy_3 + [y_1y_2y_3 \operatorname{sen}(y_1 + y_2 + y_3) - y_3 + y_1y_3\cos(y_1 - y_2 - y_3)]dy_2dy_3$
6. $\varphi^*\omega = 0$
7. $\varphi^*\omega = 16y_1^3y_2^3y_3^3y_4^3dy_1dy_2dy_3dy_4$
8. $\varphi^*\omega = dy_1dy_2$
9. $\varphi^*\omega = y_1^3y_2 \operatorname{sen}(y_3^2y_4^2)dy_2dy_3dy_4 + [y_1^2y_2^2 \operatorname{sen}(y_3^2y_4^2) - y_1y_2y_4]dy_1dy_3dy_4$
10. $\varphi^*\omega = y_1^2y_2^2y_3^2y_4^2dy_1dy_2dy_3dy_4$
11. **a.** $\varphi^*\omega = (y_1^2y_2 + y_1y_2^2 + y_2^2)dy_1dy_2$
b. $\psi^*(\varphi^*\omega) = (z_1^2z_2^2z_3z_4 + z_1z_2z_3^2z_4^2 + z_3^2z_4^2)(z_1z_3dz_2dz_4 + z_1z_4dz_2dz_3 + z_2z_3dz_1dz_4 + z_2z_4dz_1dz_3)$

Capítulo 10, Sección 4 (página 982)

2. 3/2
3. 17/2
4. 3/2
5. 0
6. 4/9
7. −7/2
8. 1/48
9. −1/48
10. 0

Capítulo 10, Sección 5 (página 991)

2. −12
3. −38
4. −51/2
5. −63/5
6. −19/2
7. −113/6
8. −182/3
9. 1115/12
10. 133/2

Bibliografía

. . . y encendió la luz para contar los pescaditos de oro que guardaba en un tarro de lata. Había diecisiete. Desde que decidió no venderlos, seguía fabricando dos pescaditos al día, y cuando completaba veinticinco volvía a fundirlos en el crisol para empezar a hacerlos de nuevo.

Gabriel García Márquez.
(Cien años de soledad)

[Ap] Tom M. Apostol, *Mathematical Analysis*, Addison-Wesley Publishing Co., Inc., 1960.

[ApI] Tom M. Apostol, *Calculus. Vol. 1*, 2a. edición, Ed. Reverté, 1982.

[ApII] Tom M. Apostol, *Calculus. Vol. 2*, 2a. edición, Ed. Reverté, 1982.

[B-F] B. M. Budak, S. V. Fomin, *Multiple Integrals, Field Theory and Series. An Advanced Course in Higher Mathematics*, Mir Publishers, Moscow, 1973.

[B-L] Peter Baxandal & Hans Liebeck, *Vector Calculus*, Oxford University Press (Oxford Applied Mathematics and Computing Science Series,) 1986.

[Bak] N. S. Bakhvalov, *Numerical Methods*, Mir Publishers, Moscow, 1977.

[Ber] G. N. Berman, *A Problem Book in Mathematical Analysis*, Mir Publishers, Moscow, 1977.

[Bu] Buck, *Advanced Calculus*, Mc.Graw-Hill Kogakusha, 3rd. edition, Ltd. 1978.

[Bu-F] R. L. Burden & J. D. Faires, *Análisis Numérico*, Grupo Editorial Iberoamérica, México, 1985.

[Br] David M. Bressond, *Second Year Calculus*, Springer-Verlag (UTM), 1991.

[C-J] R. Courant & F. John, *Introducción al Cálculo y al Análisis Matemático. Vol. 2*, Editorial Limusa, México, 1979.

[C-Sz] Lawrence J. Corwin & Robert H. Szczazarba, *Calculus in Vector Spaces*, Marcel Dekker, Inc. 1979.

[Ch-C] S. C. Chapra & R. P. Canale, *Métodos Numéricos para Ingenieros*, McGraw-Hill, 1988.

[Cle] J. M. Clement Casado, *Problemas de Cálculo Vectorial*, Editorial Alhambra, 1971.

[Cur] Caspar R. Curjel, *Exercises in Multivariable and Vector Calculus*, McGraw-Hill, 1990.

[DemI] *Problemas y Ejercicios de Análisis Matemático*, Revisado por el Profesor B. Demidovich, Editorial Mir, Moscú, 1977.

[DemII] B. P. Demidovich, *5000 Problemas de Análisis Matemático*, Paraninfo, Madrid, 1989.

[Ed] C. H. Edwards, Jr., *Advanced Calculus of Several Variables*, Academic Press, 1973.

[El] Elon Lages Lima, *Análise no Espaço* $\mathbb{R}^n$, Editôra Universidade de Brasília, 1970.

[ELI] Elon Lages Lima, *Curso de Análise. Vol. 1*, Instituto de Matemática Pura e Aplicada, Projeto Euclides, Río de Janeiro, 1981.

[ELII] Elon Lages Lima, *Curso de Análise. Vol. 2*, Instituto de Matemática Pura e Aplicada, Projeto Euclides, Río de Janeiro, 1981.

[F-K] F. Flanigan & J. Kazdan, *Calculus Two. Linear and Nonlinear Functions*, Springer-Verlag (UTM), 1990.

[Fed] *Problemas de Geometría Diferencial*, dirigido por A. S. Fedenko, Editorial Mir, Moscú, 1981.

[Fl] W. Fleming, *Functions of Several Variables*, Springer-Verlag (UTM), 2nd. edition, 1977.

[Fla] Harley Flanders, *Differential Forms, with Applications to the Physical Sciences*, Academic Press, 1963.

[Fu] Watson Fulks, *Cálculo Avanzado*, Editorial Limusa, México, 1986.

[Gr] Alfred Gray, *Tubes*, Addison-Wesley, 1990.

[Gre] *The Mathematics of Surfaces*, Editado por J. A. Gregory, The Institute of Mathematics and its Applications, Conference Series.

[H-L-S] N. Hasser, J. La Salle, J. Sullivan, *Análisis Matemático. Curso Intermedio. Vol. 2*, Editorial Trillas (Biblioteca de Matemática Superior), México, 1992.

[Ht] W. H. Hayt, Jr., *Engineering Electromagnetics*, McGraw-Hill, 5th. edition, 1989.

[KlI] Morris Kline, *Mathematics, the Loss of Certainty*, Oxford University Press, 1980.

[KlII] Morris Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, 1972.

[La] Serge Lang, *Calculus of Several variables*, Addison-Wesley, 1973.

[Li] Martin Lipschutz, *Differential Geometry*, McGraw-Hill (Schaum's Outline Series), 1969.

[M-T] J. E. Marsden, A. J. Tromba, *Cálculo Vectorial*, 3a. edición, Addison-Wesley Iberoamericana, 1991.

[M-T-W] J. E. Marsden, A. J. Tromba, A. Weinstein, *Basic Multivariable Calculus*, Springer-Verlag, 1993.

[Man] Manfredo do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, 1976.

[O'N] Barrett O'Neill, *Elementos de Geometría Diferencial*, Editorial Limusa-Wiley, S.A. México, 1972.

[PiI] Claudio Pita Ruiz, *Algebra Lineal*, Mc.Graw-Hill, México, 1991.

[PiII] Claudio Pita Ruiz, *Ecuaciones Diferenciales. Una Introducción con Aplicaciones*, Editorial Limusa, México, 1989.

[PiIII] Claudio Pita Ruiz, *Curvas Paralelas*, Miscelánea Matemática 21, 29–52 (1994).

[Sc] M. Schreiber, *Differential Forms. A Heuristic Introduction*, Springer-Verlag (Universitext), 1977.

[Sch] H. M. Schey, *Div, Grad, Curl, and all that. An Informal Text on Vector Calculus*, W. W. Norton & Company, 1973.

[SpI] Michael Spivak, *A Comprehensive Introduction to Differential Geometry, Vol. 1*, Publish or Perish, Inc. 1970.

[SpII] Michael Spivak, *Calculus on Manifolds*, W. A. Benjamin, Inc. 1965.

[W-B-F] W. Wooton, E. Beckenbach, F. Fleming, *Geometría Analítica Moderna*, Publicaciones Cultural, México, 1985.

Índice